



Human Gait Recognition under Neutral and non-Neutral Gait Sequences

By

Azhin Tahir Sabir

Department of Applied Computing

University of Buckingham

United Kingdom

A thesis submitted for the degree of Doctor of philosophy in computer science to the school of science in the University of Buckingham.

December 2015

Abstract

Rapid advances in biometrics technology makes their use for person's identity more acceptable in a variety of applications, especially in the areas of the interest in security and surveillance. The upsurge in terrorist attacks in the past few years has focused research on biometric systems that have the ability to identify individuals from a distance, and this is spearheading research interest in Gait biometric due to being unobtrusive and less dependent on high image/video quality. Gait biometric is a behavioral trait that aims to identify individuals from image sequences based on their walking style. The growing list of possible civil as well as security applications for various purposes is paralleled by the emergence of a variety of research challenges in dealing with a various external as well as internal factors influencing the performance of Gait Recognition (GR) in unconstrained recording conditions.

This thesis is concerned with Gait Recognition in unconstrained scenarios aims to address research questions covering (1) The selection of sets of features for a gait signature; (2) The effects of gender and/or recoding condition case (neutral, carrying a bag, coat wearing) on the performance of GR schemes; (3) Integrating gender and/or case classifications into GR; and (4) The role of emerging Kinect sensor technology, with its capability of sensing human skeletal features in GR and applications. Accordingly, our objectives will focus on investigating, developing and testing the performance of using a variety of gait sequence features for the various components/tasks and their integration. Our tests are based on large number of experiments based on CASIA B database as well as an in-house database of Kinect sensor recording. In all experiments, we use different dimension reduction and feature selection methods do reduce the dimensions in these proposed feature vectors, such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) and Fisher Score, followed by different classification methods like; k-nearest-neighbour (k-NN), Support Vector Machine (SVM), Naive Bayes and linear discriminant classifier (LDC), to test the performance of the proposed methods.

The initial part is focused on reviewing existing background removal for indoor and outdoor scenarios and developing more efficient versions primarily by adopting the work for wavelet domain rather than the traditional spatial domain based schemes. These include motion detection by frame differencing and Mixture of Gaussians, the latter being more reliable for outdoor scenarios. Subsequently, we investigated a variety of features that can be extracted

from various subbands of wavelet-decomposed frames of different body parts (partitioned according to the golden ratio). We gradually built sets of features, together with their fused combinations, that can be categorized as hybrid of model-based and motion-based models. The first list of features developed to deal with Neutral Gait Recognition (NGR) includes: Spatio-Temporal Model (STM), Legs Motion Detection Feature (LMD), and the Statistical model of the approximation LL-wavelet subband images (AWM). We shall demonstrate that fusing these features achieves accuracy of 97%, which is comparable to the state of the art. These features will be shown to achieve 96% accuracy in gender classification (GC), and we shall establish that the NGR2 scheme that integrates GC into NGR improves the accuracy by a noticeable percentage.

Testing the performance of these NGR schemes in recognising non-neutral cases revealed the challenges of Unrestricted Gait Recognition (UGR). The second part of the thesis is focused on developing UGR schemes.

For this, first a new statistical wavelet feature set extracted from high frequency subbands, called Detail coefficients Wavelet Model (DWM) was added to the previous list. Using different combinations of these schemes, will be shown to significantly improve the performance for non-neutral gait cases, but to less extent in the coat wearing case. We then develop a Gait Sequence Case Detection (GSCD) which has excellent performance. We will show that integrating GSCD and GC together into UGR improves the performance for all cases. We shall also investigate the different UGS scheme that generalizes existing work on Gait Energy and Gait Entropy images (GEI and GEnI) features but in the wavelet domain and in different body parts. Testing these two schemes, and their fusion, post the PCA dimension reduction yield much improved accuracy for the non-neutral cases compared to existing scheme GEI and GEnI schemes, but are significantly outperformed by the last scheme. However, by fusing the UGS scheme with the GSCD+GC+UGR scheme above we will get best accuracy that outperform the state of the art in GR specially in the non-neutral cases.

The thesis ended by conducting a rather limited investigation on the use of the Kinect sensors for GR. We develop two sets of features: Horizontal Distance Features and Vertical Distance Features from small set of skeleton point trajectories. The experimental result on neutral was very successful but for the unrestricted gait recognition (with the 5 case variations) satisfactory but not optimal performance relies on the gallery including balanced number of samples from all cases.

Acknowledgments

Allah the Most Gracious and Merciful: Who gave me energy, health, nerves and provided me with all the people to whom I am dedicating this hard work, which took a lot of determination and time until it came to light.

My parents: I would like to designate the entire fruitful outcome of this work to my father who has been waiting for so long to see the result of his son's work, to my mother, who hasn't stopped praying for this work to be done.

My Family: I cannot put my thankfulness and appreciation to my wife **Shiraz** for her patience and endless love, who has stood by me through all my travails, my absences, my fits of pique and impatience. She gave me support and help, discussed ideas and prevented several wrong turns. She also supported the family during much of my graduate studies. Along with her, I want to acknowledge my two sons, **Allen** and **Alvand**, they have never known their dad as anything but a student, it seems. Good boys, both, and great sources of love and relief from scholarly endeavor.

My Supervisors: My greatest thanks go to my supervisors Prof. Sabah Jassim and Dr. Naseer Al-Jawad., they continuously support me during the development of this thesis and they always encouraged and motivated me to continue even in difficult times. Without their viable advice, guidance, and ideas, this thesis would have never been come to light.

My thankfulness and appreciation to my extended family, particularly brothers, and sisters, for their support, encouragement, patience and praying for me during the time of my study.

My Friends: Last but not least, many thanks are due to all my friends, fellow students, and all members of staff at the applied computing department, University of Buckingham for the fruitful collaboration work and for all their support and important discussions.

Table of Contents

Abstract.....	i
Acknowledgments.....	iii
Table of Contents.....	iv
List of Figures	viii
List of Tables	x
Abbreviations.....	xii
Declaration.....	xv
 Chapter One Introduction	1
1.1 Automatic human identification.....	1
1.2 Biometric Systems.....	2
1.2.1 Factors influence choice of biometric traits.....	4
1.2.2 Biometric System principle working processes	6
1.3 Popular Biometrics Traits.....	7
1.3.1 Voice recognition.....	7
1.3.2 Face recognition.....	8
1.3.3 Gait recognition	8
1.4 The scope of research in this thesis	9
1.5 Gait recognition challenges	9
1.6 Thesis Contribution	10
1.7 List of Publications.....	11
1.8 Thesis Organization.....	11
 Chapter Two Background and Literature survey	12
2.1 Pre-processing	12

2.1.1	Background Subtraction techniques	13
2.1.2	Gait Cycle Estimation	16
2.2	Gait Feature extraction and Classification	18
2.3	Human gait recognition using conventional camera	18
2.4	Human gait recognition using Kinect sensor	25
2.5	Gender classification based on gait features	26
2.6	Chapter Summary.....	28
Chapter Three Human gait Recognition System		29
3.1	Pre-processing	29
3.1.1	Background Subtraction technique	30
3.1.2	Gait Cycle estimation and Normalisation	37
3.2	Feature Extraction	37
3.2.1	Spatio-temporal Model	38
3.2.2	Legs Motion Detection	40
3.2.3	Approximation and Detail coefficients Wavelet Model	41
3.2.4	Gait Energy Image	43
3.2.5	Gait Entropy Image.....	44
3.3	Dimension reduction	46
3.3.1	Feature Transformation.....	47
3.3.2	Feature Selection.....	49
3.4	Classification Methods.....	50
3.4.1	k-Nearest-Neighbour.....	50
3.4.2	Support Vector Machine	51
3.4.3	Linear discriminant classifier.....	52
3.5	Databases.....	53
3.6	Chapter Summary.....	53
Chapter Four Neutral gait sequences		54

4.1	Person Recognition from Neutral Gait Sequences	54
4.1.1	The Neutral Gait Recognition (NGR).....	54
4.1.2	The Experimental dataset - CASIA Database.....	56
4.1.3	Experiment Setup.....	56
4.1.4	Experimental Results	57
4.2	The Reliability of NGR decisions	62
4.3	Gender classification from neutral gait sequences	64
4.3.1	Gender Classification Scheme (GCS).....	64
4.3.2	Experiment Setup.....	65
4.3.3	Experimental Results	66
4.3.4	Reliability of GCS Decisions.....	69
4.4	The 2-level Neutral Gait Recognition (NGR2)	70
4.4.1	The NGR2 Algorithm	70
4.4.2	Experimental Results	71
4.5	Conclusions	72
	Chapter Five Unrestricted Gait Recognition	74
5.1	Extending NGR to Unrestricted Gait Recognition.....	74
5.1.1	The modified Unrestricted Gait Recognition (UGR) algorithm.	77
5.1.2	Experimental Results	77
5.2	Gait Sequence Case Detection (GSCD)	79
5.3	The GSCD-UGR scheme	80
5.4	Unrestricted Gender Classification (UGC)	82
5.4.1	The proposed UGC method	82
5.4.2	Experimental Results	84
5.5	Incorporating Gender classification into GSCD-UGR.....	85
5.6	Unrestricted Gait Schemes using Gait Energy and Entropy Image (UGS).....	87
5.6.1	UGS-(GEI & GEnI)	87

5.6.2	Description of Experiments	89
5.6.3	Experimental Results	89
5.7	Combining UGS-(GEI & GEnI) with UGR.....	91
5.8	Comparison of our methods with the State Of The Art	92
5.9	Conclusions	94
	Chapter Six Gait recognition using Kinect sensor	96
6.1	Human Skeleton Tracker.....	96
6.2	Kinect Database.....	98
6.3	Gait Recognition Using Kinect Sensor	99
6.3.1	Feature extraction based on Kinect.....	99
6.3.2	Performance of the Kinect Neutral Gait Recognition (KNGR).....	102
6.3.3	Unrestricted Gait Recognition Using Kinect Sensor (KUGR)	104
6.4	Conclusions	109
	Chapter Seven Conclusions and Future Research.....	110
7.1	Conclusions	110
7.2	Future Work	114
	References.....	115

List of Figures

Figure 1-1: Examples of biometric characteristics .	4
Figure 1-2: Biometric system working processes diagram	7
Figure 2-1: The complete gait cycle: stance and swing	16
Figure 2-2: Gait Cycle estimated based on 4 frames which represent one gait cycle	18
Figure 2-3: Gait signature by creating a triangles from hands and feet	20
Figure 2-4: Gait signature using static, activity-specific parameters.	20
Figure 2-5: Dividing the silhouette into 7 regions.	27
Figure 3-1: Background subtraction Motion Compensation	31
Figure 3-2: Background Subtraction based on MoG.	32
Figure 3-3: Background subtraction in the wavelet domain.	34
Figure 3-4: Gait cycle estimation overview	37
Figure 3-5: Spatio-temporal Model (STM) feature vector.	38
Figure 3-6: Golden ratio proportions for human body	39
Figure 3-7: Human body when the legs are overlapping or wide apart.	40
Figure 3-8: The overview of Legs Motion Detection (LMD) feature vector.	41
Figure 3-9: AWM example (a) Lower human body part (b) Upper human body	43
Figure 3-10: Examples of Gait Energy Image.	43
Figure 3-11: Examples of Gait Entropy Image.	45
Figure 3-12: Overview of the skeleton points provided by Kinect Sensor.	46
Figure 3-13: Support Vector Machine Overview	51
Figure 3-14: Linear Discriminant Classifier Overview.	53
Figure4-1: Overview of the Neutral Gait Recognition (NGR) system.	55
Figure 4-2: The Recognition Performance of NGR using different feature vectors.	58
Figure 4-3: The Recognition Performance of NGR with LDA.	59
Figure 4-4: Recognition Performance of NGR based on fusing the three feature sets (LMD, STM and AWM) with different number of dimensions using LDA.	61
Figure 4-5: Overview of the reliability in the proposed NGR method.	62
Figure 4-6: Overview of Gender Classification Scheme based on neutral gait sequences.	65
Figure 4-7: The Average and standard deviation of the GCS. using k-NN (k=1).	67
Figure 4-8: The Average and standard deviation of the GCS using SVM.	67
Figure 4-9: Performance Comparison result of using k-NN and SVM.	68

Figure 4-10: The effect of varying the number of folds in the cross validation.	68
Figure 4-11: Overview model for combining gender classification with gait recognition.....	71
Figure 5-1: Recognition Performance of using NGR feature vectors for UGR. ..	75
Figure 5-2: Recognition Performance of the UGR for S1 and S2 feature vectors separately.	78
Figure 5-3: Recognition Performance of Gait Sequence Case Detection.....	80
Figure 5-4: Recognition Performance of incorporating UGR with GSCD.....	81
Figure 5-5: Recognition Performance of GSCD-UGR using different number of samples ...	81
Figure 5-6: The percentage of female candidates that misrecognized and confused as a males in CW gait sequences, using UGR.....	82
Figure 5-7: Generating gallery and probe set from the CASIA B gait database for gender classification based on unrestricted gait sequences.	83
Figure 5-8: Recognition Performance of Unrestricted Gender Classification (UGC)	84
Figure 5-9: Overview model combining GSCD, UGC, and UGR	86
Figure 5-10: Unrestricted Gait Schemes (UGS) using Gait energy and Gait Entropy	88
Figure 5-11: Recognition Performance of UGS based on WGEI and WGE _n I using PCA and LDA separately.....	89
Figure 6-1: Kinect database locations.....	98
Figure 6-2: Overview of gait recognition system using Kinect Sensor	99
Figure 6-3: Overview of the skeleton points from Kinect Sensor.	100
Figure 6-4: Recognition Performance of KNGR using Kinect database1.....	103
Figure 6-5: Kinect database-2 example.	105
Figure 6-6: Recognition Performance of KUGR based on first scenario.	106
Figure 6-7: Recognition Performance of KUGR based on second scenario.	107
Figure 6-8: Recognition Performanc of KUGR based on third scenario.....	107
Figure 6-9: Recognition Performance of KUGR based on fourth scenario.....	108
Figure 6-10: Recognition Performance of KUGR based on fifth scenario.....	108

List of Tables

Table 1-1: Comparison of Various Biometric Technologies. High, Medium, And Low Are Denoted By H, M, and L, Respectively	5
Table 3-1: Background subtraction- Motion Compensation process time using four video sequences taken from our own indoor video sequences.	31
Table 3-2: Background Subtraction –MoG, process time based on action database.....	33
Table 3-3: MoG in the wavelet domain, process time using the combination of the three non-approximation coefficients (LH, LV and HH), based on action database.	35
Table 3-4: Refined MoG in the wavelet domain process time, using Vertical coefficient, based on action database.	36
Table 4-1: Recognition Performance based on NGR (%) using different number of samples in the gallery set.....	60
Table 4-2: Performance Comparison between proposed NGR method with other methods in the literature using leave- one- out cross validation, (based on CASIA B gait database)	61
Table 4-3: Recognition Performance of NGR and the reliability of its decisions	63
Table 4-4: The percentage of confusing female with male and also male with female using Neutral Gait Recognition (NGR).....	64
Table 4-5: The performance of Gender Classification Scheme using kNN, k=1, 3, and 5	67
Table 4-6: Performance Comparison of GCS with other methods using CASIA B Gait database, CR is Classification Rate.	69
Table 4-7: The percentage of confusing female with male in Neutral Gait Recognition (NGR) based on FV1 and FV2.	70
Table 4-8: Recognition Performance of NGR with and without gender classification.	71
Table 5-1: Recognition Performance of using UGR for a limited number of combinations of gallery to probe ratios using S2 {STM+LMD+ lower body part (AWM+DWM)) feature	79
Table 5-2: Recognition Performance of Unrestricted Gender Classification (UGC) for each feature vector (STM, AWM, and DWM) separately and their fusion schemes.	84
Table 5-3: Performance Comparison of UGC with results published in the literature Using CASIA B database	85
Table 5-4: Recognition Performance of UGR based on CW, with and without UGC. Using CASIA B database.	86

Table 5-5: Recognition Performance of Unrestricted Gait Schemes (UGS) based on GEI, GEnI, and different combining between them.....	90
Table 5-6: Recognition Performance of UGS based on fusing GEI and GEnI with each of WBP, LBP, and UBP and different combinations between them	91
Table 5-7: Recognition Performance of UGS based on combination of GEI and GEnI (WBP+LBP) with different gait sequences in the gallery and probe set..	91
Table 5-8: Recognition Performance of gait recognition based on fusing UGS and UGR using score level fusion.	92
Table 5-9: Recognition Performance of gait recognition based on fusing UGS and UGR using decision level fusion.	92
Table 5-10: Performance Comparision between our methods and the State of the Art – standard protocol..	93
Table 5-11: Performance Comparision between our methods and the State of the Art – mixed protocol.	93
Table 6-1: Joint points name and numbers from Kinect Sensor (N is number).	97
Table 6-2: Performance Comparison according to recognition rate between proposed method and other methods in the litriture.	104
Table 6-3: Gallery and probe samples set in five scenarios.....	105

Abbreviations

ACDA	Adaptive Component and Discriminant Analysis
ANN	Adaptive Neural Network
AVG	Average Different Sequence
AWM	Approximation coefficients Wavelet Model
BGEI	Backfilled Gait Energy Image
CASIA	Chinese Academy of Sciences, Institute of Automation
CB	Carrying Bag sequences
CBB	Carrying Bag on Back gait sequences
CBS	Carrying Bag over Shoulder gait sequences
CR	Classification Rate
CS-LBP	Conditional-Sorting Local Binary Pattern
CV	Cross Validation
CW	Coat Wearing sequences
DRP	Distance of Recognized Person
DWM	Detail coefficients Wavelet Model
DTW	Dynamic Time Warping
2D-DCT	2-Dimensional Discrete Cosine Transform
DWT	Discrete Wavelet Transform
EER	Equal Error Rate
EHMM	Embedded Hidden Markov Model
FA	False Acceptance
FDHI	Frame Difference History Image
FGEnI	Feature selection based on Gait Entropy Image
FLHD	Frontal Leg Hamming Distance
FPCC	Frontal Legs Pixel Change Counter
FR	Frame Reference
FS	Feature Selection
GC	Gender Classification
GCS	Gender classification Scheme
GEI	Gait Energy Image
GEEnI	Gait Entropy Image
GHEI	Gradient Histogram Energy Image
GPPE	Gait Pal and Pal Entropy
GR	Gait Recognition
GMM-HMMs	Gaussian Mixture Model-Hidden Markov Models
GSCD	Gait Sequence Case Detection
HR	Highly Reliable
HCR	High level confident
HDF	Horizontal Distance Feature
HMM	Hidden Markov Model
HOF	Histogram of Optical Flow

HOG	Histogram of Image Gradient
HR	Highly Reliable
iHMM	individual HMM
KNGR	Kinect Neutral Gait Recognition
k-NN	k-Nearest Neighbor
KUGR	Kinect Unrestricted Gait Recognition
LBP GEI	Lower Body Part GEI
LDA	Linear Discriminant Analysis
LDC	Linear Discriminant Classifier
LF	Local Binary Pattern flow
LMD	Legs Motion Detection
MoG	Mixtures of Gaussians
NB	Naive Bayes
NGR	Neutral Gait Recognition
NGR2	2-level Neutral Gait Recognition
Nu	Neutral gait sequences
o HMM	offline HMM
OCM	Outclass Members
ORR	Outclass Rejection Rate.
PCA	Principal Component Analysis
PCs	Principal Components
PRWGEI	Poisson Random Walk of Gait Energy Image
PRWGEI	Poisson Random Walk of Gait Energy Image
RDWT	Redundant Discrete Wavelet Transform
RLHD	Rear Leg Hamming Distance
RPCC	Rear Legs Pixel Change Counter
RR	Recognition Rate.
RS	Recognition Strength
RP	Recognized Person
SDK	Software Development Kit
SBW	Silhouette-Binary-Wavelet
SEW	Silhouette-Edge-Binary Wavelet
SGEI	Structural Gait Energy Image.
SGW	Silhouette-Grey-Wavelet
ShGEnI	Shannon entropy
SSW	Silhouette Skeleton Wavelet
Std	standard deviation
STIPs	space-time interest points method
STM	Spatio-temporal Model
SVM	Support Vector Machine
TR	True Rejection
UBP	Upper Body Part
UGC	Unrestricted Gender Classification
UGR	Unrestricted Gait Recognition

UGS	Unrestricted Gait Schemes
USF	University of South Florida
VDF	Vertical Distance Feature
WCM	Within Class Member
WLC	Wearing Long Coat gait sequences
WMS	Misrecognized rate
WSC	Wearing Short Coat gait sequences
WT	Wavelet Transform

Declaration

I hereby declare that all the work in my thesis entitled “**Human Gait Recognition under Neutral and non-Neutral Gait Sequences**” is my own work except where due reference is made within the text of the thesis.

I also declare that, to the best of my knowledge, none of the material has ever previously been submitted for a degree in the University of Buckingham or any other university.

Chapter One

Introduction

Biometrics refers to the automated use of behavioural or physiological characteristics to determine or verify the identity of a person. Terrorist attacks in the past few years have led to increased interest in identifying/verifying persons at a distance. Terrorists and other criminals often are caught on CCTV cameras from a relatively far distance, and the sooner they are detected, the quicker the response need to be made to prevent their intended crime. Gait is one of the few biometrics that can be detected unobtrusively at a distance, and is difficult to conceal. Extracting gait signature and recognising the person from low resolution images become a major challenge for researchers. The performance of gait biometric systems is influenced by a variety of covariate factors beyond the low resolution images at a distance. Some of these factors are external like type of worn clothes, carrying bags as well as internal factors such as pregnancy and overweight. This thesis is an attempt to develop and test the performance of gait recognition schemes that could be used for person's identification from a distance taking the account of the above external factors. We label external factors as non-neutral if the person is carrying bags, or wearing coats, while Neutral conditions means that the person is not carrying objects or wearing long coats.

In the rest of this chapter we shall start by providing general explanations and basic concepts for biometric systems, popular biometrics traits, highlight gait recognition (Importance, Challenges and Applications), problem description, and end with thesis objectives and contributions together with thesis organization.

1.1 Automatic human identification

The importance of biometric systems in today's world has been supported by the need for large scale identity management systems whose functionality relies on the accurate determination of an individual's identity (Katiyar, et al., 2013). This has been especially the case in the past ten years, in which terrorist attacks have happened frequently, which have caused people to focus on the importance of security monitoring and control in national defence and public safety. The identification of a person is becoming increasingly important as multiple numbers of ID cards, punches,

secret passwords and PINs that are used for personal identification. The aforementioned biometrics or characteristics are tightly related to an individual and cannot be forgotten, shared, stolen, used by anyone else or easily hacked.

Many biometric technologies have emerged for identifying or verifying individuals by analysing face, fingerprint, palmprint, iris, gait or a combination of these traits. Human recognition methods that are image-based like fingerprints, palmprint or iris biometric modalities, usually require a cooperative subject and physical contact or close nearness. In monitoring scenes, people are usually away from the cameras, which make most biometric features no longer available. In fact, many of the above biometric methods are not suitable to reliably recognize non-cooperating individuals at a distance in the real world. Gait that aim to recognize human by the way they walk can be used without having the above mentioned disadvantages. Using gait as a behavioural biometric system is strongly motivated by the need for an automated recognition system for visual surveillance and monitoring applications (Wang, et al., 2003).

1.2 Biometric Systems

Biometrics are defined as measurable physiological and/or behavioural characteristics, Physiological biometrics examines physiological characteristics such as: iris, faces, fingerprints, DNA, and hand geometry, while behavioural biometrics examines behavioural issues such as signature and gait (Wayman, et al., 2005), (see Figure 1.1). The term biometric comes from the Greek words bios (life), and metrikos (measure) (AlMahafzah & AlRwashdeh, 2012). The history of biometrics includes the identification of individuals by unique bodily features, the first evidence of biometrics appeared in 29,000BC when cavemen used their fingerprints to sign their drawings (Idrus, et al., 2013). Another early use of biometrics includes the practice in ancient China whereby babies were distinguished from each other through ink stamps of palmprints and footprints (Uhl & Wild, 2008). In the early nineteenth century, biometrics was mainly used in criminology, when researchers investigated the relationship between physical features and criminal tendencies (Sentayehu, 2006).

Biometric systems can act in one of two different modes and functionality: an 'identification' mode or a 'verification' (authentication) mode. An identification system can be defined as one which carries out processes to determine a person's identity by

performing matches against multiple biometric templates. Identification systems can be categorised into two types: positive and negative identification. Positive identification systems verify a positive claim of enrolment which is designed to find a match for a user's biometric information in a database of biometric information. Basically, this type of identification will answer the question: "WHO AM I?" A positive identification model system would be a prison release program where users do not enter an ID number or use a card, but simply gait signature extracted from video sequences (as an example of biometric system) and are identified from an inmate database (Wayman, et al., 2005). Negative identification systems search databases in the same way, matching one template against many, however are designed to ensure that a person is not present in a database. This prohibit individuals from enrolling more than one time in a system, and is frequently used in large-scale public benefits programs in which users were enrolled multiple times to gain benefits under different names. Negative identification applications cannot be made voluntarily. Negative identification can be found in applications such as driver licensing and social service eligibility systems where multiple enrolments are illegal and a user claims not to be previously enrolled (Wayman, et al., 2005). Fingerprinting and retinal scanning applications are two types of biometric systems that are used in documented negative identification. Not all identification systems are based on determining a username or ID, some systems are designed to determine if a user is a member of a particular category. For example, an airport may have a database of known terrorists with no knowledge of their real identities. In this situation the system would return a match, without having knowledge about person's identity is involved.

Verification systems seek to answer the question "Is this person who they say they are?" based upon verifying a sample collected versus a previously collected biometric sample for the individual. When biometrics are implemented for verification purposes, they will answer the question: "AM I WHO I SAY I AM?" while an identification system looking for the question "Who am I?" (Gafurov, 2007). Certain verification systems perform very limited searches against multiple enrolled records. For example, an individual with three enrolled fingerprint templates may be able to place any of the three fingers to verify, and the system achieves 1:1 matches against the user's enrolled templates until a match is found. If you are considering getting either one of these biometrics identity verification systems, it is important to know the differences

between them, because the biometrics technologies for both are advancing continuously.

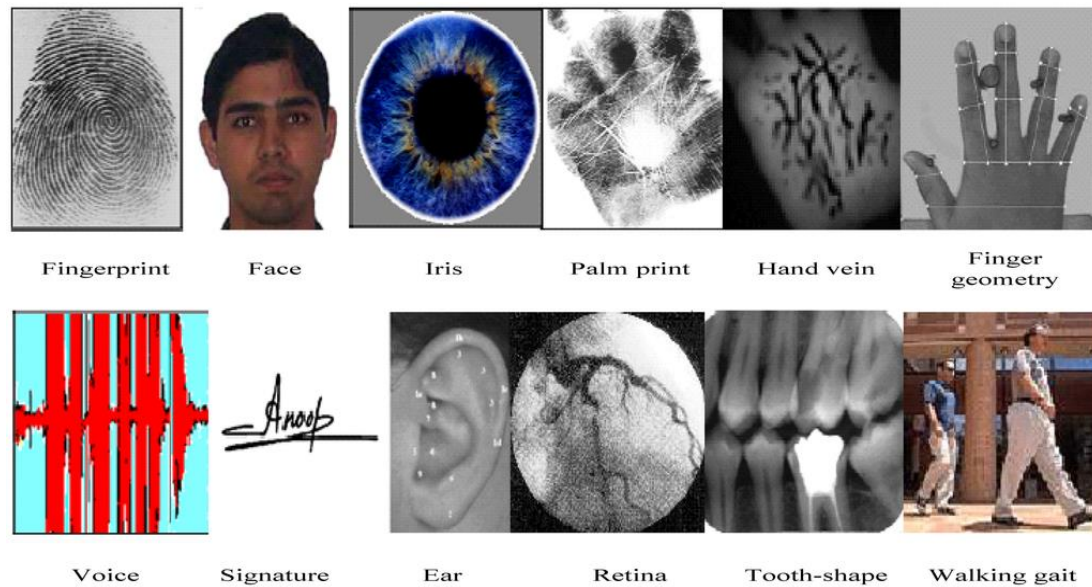


Figure 1-1: Examples of biometric characteristics (Zhao, et al., 2013).

1.2.1 Factors influence choice of biometric traits

Automatic biometric systems have been presented over the last few decades, due to scientific progress in computer processing field (Bobde & Satange, 2013). There are numerous biometrics in use today and a variety of biometrics that are still in the early stages of development. Biometrics can be classified into two groups: those that are currently in use (fingerprint, face, iris and voice) and others that are still being researched or are under expansion, like Earshape and gait (Mir, et al., 2011). There are several factors that shows the necessity of each type in the biometric systems, Table 1.1 shows the comparison of these types of biometric system including Universality, Uniqueness, Collectability, Permanence, Performance, Acceptability and Circumvention. For example, it is far-famed that fingerprint-based methods are more accurate than the voice-based methods. Conversely, in a tele-banking application, in a reverse the voice-based technique may be preferred (Jain, et al., 2004).

Currently, biometric techniques are used mainly in security operations. For example, they are used in secure visitor systems, state benefit payment systems, border controls and security surveillance. Surveillance technology is now pervasive in modern society; this is ascribable to the increasing number of crimes as well as the vital need

to provide a more secure environment. Owing to overgrowth in the number of security cameras and of the need for enough manpower to supervise them, the deployment of non-invasive biometric technologies becomes important for the development of automated visual surveillance systems and forensic investigation. In the last few years there has been a large number of papers introducing new systems, like Earshape, keystroke dynamic, and gait. Recently, the use of gait for person's identification in surveillance applications has attracted researchers from the computer vision community (Bouchrika, et al., 2011). If gait-recognition matured, it could well be used for police surveillance applications ((PITO), 2005) . Perceiving gait from a distance makes gait convenient for use in surveillance systems.

Table 1-1: Comparison of Various Biometric Technologies. High, Medium, And Low Are Denoted By H, M, and L, Respectively (Jain, et al., 2004).

Biometric identifier	Universally	Distinctiveness	Permanence	Collectability	Performance	Acceptability	Circumvention
DNA	H	H	H	L	H	L	L
Ear	M	M	H	M	M	H	M
Face	H	L	M	H	L	H	H
Facial thermo gram	H	H	L	H	M	H	L
Fingerprint	M	H	H	M	H	M	M
Gait	M	L	L	H	L	H	M
Hand geometry	M	M	M	H	M	M	M
Hand Vein	M	M	M	M	M	M	L
Iris	H	H	H	M	H	L	L
Keystroke	L	L	L	M	L	M	M
Odor	H	H	H	L	L	M	L
Palmprint	M	H	H	M	H	M	M
Retina	H	H	M	L	H	L	L
Signature	L	L	L	H	L	H	H
Voice	M	L	L	M	L	H	H

1.2.2 Biometric System principle working processes

In general biometric systems have unique principles during the time it is conducted. To present these principles we explain each one separately as follows:

- i. **Presentation:** the process of biometric data gathering by which the user presents his/her biometric data to specific devices, such as placing a finger on the plate of a finger reader device, or camera that record the person gait sequences.
- ii. **Feature Extraction:** is the process of defining a set of characteristics (features), which will most efficiently represent the information that is important for analysis and classification. Feature extraction takes place during enrolment and verification, any time a template is generated.
- iii. **Biometric Data:** unprocessed image data presented by the user which is also referred to the raw biometric data or biometric sample, it is used to generate a biometric template with the help of feature extraction processes.
- iv. **Template:** a digital reference of distinct characteristics that have been extracted from a biometric sample. A template size can vary in size between a few bytes for hand geometry to several thousand bytes for facial or gait recognition.
- v. **Enrolment:** the individual provides a sample of the biometric which is captured by a device (e.g. a camera). Information is extracted from this sample to create a biometric representation (template) which is stored in the database.
- vi. **Matching:** A process where a stored template is compared with a live template at the time of verification and a decision made, based on the score gained, that a user is the authenticated human or not. In some applications, certain identity attributes about the person (name, ID number, etc.) is also stored along with the biometric reference (Srivastava, 2013) (see Figure 1.2).

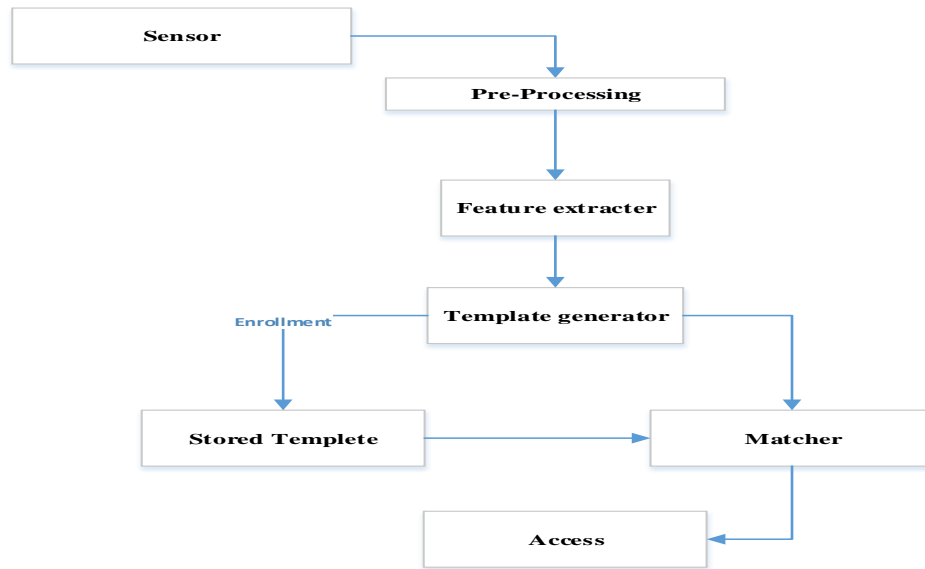


Figure 1-2: Biometric system working processes diagram

1.3 Popular Biometrics Traits

Generally based on different types of biometric modalities, two basic branches can be identified: physiological (or passive) and behavioural (or active). In this approach we can present another classification based on the way of features gathering into; those that need physical contact like (fingerprint, hand geometry, and Palmprint) and the other that can be captured from a distance like (face, voice and gait). Identifying human at a distance is in rising demand in computer vision. This is due to need it for automated human identification in surveillance systems. Humans identify each other based on their various characteristics we overview some biometric system that recognize human at a distance.

1.3.1 Voice recognition

Voice Recognition is a skill which allows the user to use his/her voice as an input device, it uses a measurable, physical characteristic, or personal behavioural characteristic to verify and authenticate an individual. Voice recognition biometric systems use the distinctive aspects of the voice to verify the identity of individuals. Voice recognition is sometimes confused with speech recognition, a method which translates what a user is saying (a process unrelated to authentication). Human voice is related to, voice recognition, speech recognition and natural language processing, Voice recognition allows you to provide input to an application with your voice. The

voice recognition process is performed by a software component known as the speech recognition engine. The main function of the voice recognition engine is to process spoken input and translate it into text that an application understands (Reynolds, 2002).

1.3.2 Face recognition

Face recognition is a physiological biometric recognition system that aims to recognize individuals from a distance. In the beginning of the 1970's, face recognition took another step towards automation and was treated as a 2D pattern recognition problem. Face recognition is the relevant applications of image analysis. Generally face recognition is proposed for security issues. Face recognition relates to a defined process of the brain: although the human ability to identifying known faces is relatively good, we face difficulty when we deal with a large number of unknown faces (De Carrera, 2010). In the last decade face recognition developed from still images and video streams into an active research area and became beneficial in many biometric applications. These include information security, law enforcement, access control, smart cards and surveillance systems.

1.3.3 Gait recognition

Over the past few decades gait recognition has been investigated in different fields including medical applications, Analysis of abnormal gait signature, Bio-inspired Bipedal robot controller and finding the effect on gait with different age, body, and weight. Gait as a means of biometric recognition aims to recognize a subject by the way they walk. Psychological investigations have revealed that human are not good in recognising other than individuals that they know and even then their ability is limited but statistically significant. However, Physiological studies have also shown that human are able to recognize the gender of a walker (Boyd & Little, 2005). Medical work supports the view that if all gait movements are measured then the gait is unique (Murray, 1967). Gait recognition could be used in different applications such as identifying unauthorized persons, identifying their gender, making predictions about individual age or physical health, and determining walking-related abnormalities by analyzing the way they walk or move. In the field of automatic biometric recognition for human identification, gait recognition has advantages over other biometrics like

fingerprints, iris, face and other biometric features. These advantages includes: Gait signature can be captured at a distance, the person identified does not need to participate in the biometric feature capturing process this is called (Unobtrusive), and recognition can be performed in low resolution degraded images. Gait information is very difficult to hide, and attempting to do attracts suspicion while face recognition that might be tricked in various ways.

Applications of gait recognition

In general, gait applications can be categorised into three types: Health applications use human gait for patient walking analysis for detection of physical abnormalities (Cho, et al., 2009). Robotics applications where human gait is used to create/enhance a stable walking gait on a Biped Robots (Hafner & Bachmann, 2008). And Security applications where gait recognition can be used for identifying unauthorized intruding persons or criminals, for forensics purposes by attempting to identify a known abnormality of a suspect's movement, or detecting and recognizing a person who carries suspicious objects effecting the way he/she walks. (Savage, et al., 2013).

1.4 The scope of research in this thesis

This thesis is focused on automatic recognition of individuals from their gait while walking without restriction on their clothing conditions or carrying bags/objects. We shall conduct our investigations in accordance with standard biometric strategies and approaches in relation to feature extraction, adopted classification models, experimental protocols and standard evaluation schemes. I am primarily motivated by the growing list of potential applications as well as the relatively new research area with serious stimulating challenges. At this stage, I am more interested in the security applications related to fighting crime and terrorism.

1.5 Gait recognition challenges

There are apparent challenges in gait capturing that make it extremely difficult to identify and record all parameters that affect gait. Gait recognition problems are challenging due to largely varying appearances. The challenges are overcome gait motion variations due to numerous conditions such as, clothing, walking surface. To present these factors that may negatively affect the performance of gait recognition we

classify these factors into external and internal factors. Viewing angles, illumination changes, clothes, walking surface, types of shoes and objects carried are classified as an external factors. Sickness of gait, drunkenness, aging, pregnancy and gaining or losing weight are classified as internal factors. to overcome some of the main challenges of current gait analysis methods in this thesis we focus on external factors.

1.6 Thesis Contribution

The main aim of this thesis is to design gait recognition scheme under neutral and non-neutral gait sequences. There are different methods adopted for background subtraction in the literature. In this thesis we contribute in proposing a method based on Mixture of Gaussian (MoG) and wavelet transform. Moreover we proposed different sets of feature for gait recognition: Spatio-temporal Model (STM) is the first feature, which aim to deal with the distances between feet, hands, shoulder and height. The second feature set is Legs Motion Detection (LMD) used to detect the motion of legs during gait cycle. The third feature vector is statistical properties of wavelet transform called Approximation and Non Approximation wavelet model (AWM and NAWM respectively). We also investigated the use of the traditional features of Gait Energy Image (GEI) and the Gait Entropy Image (GEnI), to be used in new form. We have also concentrated our analysis on the human gait, trying to get the advantages that the Kinect camera provides, this is by proposing two sets of feature: Vertical Distance Feature (VDF) and Horizontal Distance Feature (HDF). To identify/classify human gait in this thesis we used various classification methods like k-NN, SVM, Naive Bayes and Linear Discriminant Classifier (LDC) Applied to three different databases, which are CASIA B gait database as a well-known indoor database; and another two different databases that we have created ourselves based on Kinect sensor. Basically, Gait Recognition (GR) and Gender Classification (GC) methods based on neutral and non-neutral gait sequence have been proposed. Additionally, the Reliability method for the proposed GR and GC for neutral gait sequences is suggested. Furthermore Incorporating GC and a new proposed method called Gait Sequence Case Detection (GSCD) with GR have been also shown to improve the performance of GR. And finally Gait Recognition using Kinect sensor based on Neutral and Non-neutral gait sequence is also investigated.

1.7 List of Publications

- **Azhin Sabir**, Naseer Al-jawad, and Sabah Jassim. "Feature selection gait-based gender classification under different circumstances". SPIE Real-Time Image and Video Processing, part of SPIE Photonics Europe, 2014.
- Mohammed H. Ahmed, Naseer Al-jawad, and **Azhin Sabir**. "Gait recognition based on Kinect sensor." SPIE Real-Time Image and Video Processing, part of SPIE Photonics Europe, 2014.
- **Azhin Sabir**, Naseer Al-jawad, and Sabah Jassim.. "Human gait gender classification based on fusing spatio-temporal and wavelet statistical features." Computer Science and Electronic Engineering Conference (CEEC), 2013 5th. IEEE, 2013.
- **Azhin Sabir**, Naseer Al-jawad, and Sabah Jassim. "Gait recognition using spatio-temporal silhouette-based features." SPIE Defense, Security, and Sensing. International Society for Optics and Photonics, 2013.

1.8 Thesis Organization

The rest of this thesis is organized as follows. Chapter 2 will provide a critical review of previous and current work relating to gait recognition and gender classification methods. In Chapter 3, we present gait recognition process steps including; object detection, feature extraction, dimension reduction and classification methods. In Chapter 4 we present our proposed methods in human gait recognition and gender classification based on neutral gait sequence, this chapter will also highlight the reliability in biometric identification/recognition and gender classification. Chapter 5 deals with gait recognition under unrestricted gait sequences: in this chapter we present two different challenges in gait recognition posed by carrying bags and wearing a coat, and also we show our proposed methods to cope with these two gait sequence variations. In this chapter we also present a new method called gait sequence case detection that aim to detect human gait sequence before the process of identification. In Chapter 6, we will focus on human gait recognition using Kinect camera: in the proposed method we use neutral and non-neutral gait sequences. Finally, the merits of this thesis and possible future research areas are pointed out in Chapter 7.

Chapter Two

Background and Literature survey

Gait recognition is a typical example of biometric recognition tasks and building a gait recognition system follow the same structure as depicted in Figure 1.2 in chapter 1. Developing such system is based on designing a digital representation of the human gait, called feature vector to be extracted from video of the walking person recorded by a camera from a distance. Recognition/identification of the person in a video will be based on measuring the similarity between the videoed person's extracted fresh gait feature vector and a set of already stored labelled gait feature vectors, called templates for people enrolled on the system. The extraction of gait feature vectors from videos of persons walking by the recording camera need to be first pre-processed to remove background regions and capture the persons gait cycles from which some numerical values can be determined automatically as attributes of the gait feature vector.

In this chapter we present background description of the various components of gait recognition and conduct a literature survey of existing research in these components. The chapter begins with the pre-processing component in section 2.1. Section 2.2 will deal with feature extraction component, while in section 2.3 we focus on gait recognition approaches in which gait sequences are captured based on conventional camera under covariate factors. Section 2.4, deals with gait recognition method where gait sequence were captured based on skeleton data using the Kinect sensor depth camera. Due to many factors including differences in body structure between male and female as well as differences in the way they wear dresses and carry items, many research work has been focused on gender recognition based on gait feature. To highlight these methods the last section will focus on human gender classification based on gait features using neutral and non-neutral gait sequences.

2.1 Pre-processing

The main aim of pre-processing for any gait recognition scheme is to segment the video frames so as to remove the background in the frames and only retain the image of the person, under consideration, as well as his/her gait cycle. Pre-processing may also include procedures to improve the video frames data quality by removing some

undesired distortions that could hinder the removal of the background and the ultimate gait image analysis. In this section we shall focus on existing research and methods the pre-processing step includes: background subtraction, normalisation and gait cycle estimation.

2.1.1 Background Subtraction techniques

During pre-processing step, background subtraction is a very common technique used in this field. There are a number of techniques that are used for this purpose, some of which work in the spatial domain (Mayo & Tapamo, 2009; Migliore, et al., 2006; Kameda & Minoh, 1996; (Heikkilä, 2004) and Zhou & Hoang, 2005) while others work in frequency domain primarily using wavelet transforms (Aghaee, et al., 2011; Liu & Wang, 2009; Chang, et al., 2007 ; Crnojevic, et al., 2009 and Tao, et al., 2009). The wavelet transform approach has two advantages in background subtraction methods; the first of which is reducing computational complexity, and secondly to reduce noise.

(Mayo & Tapamo, 2009) Proposed the use of frame differencing to detect moving objects. Calculating the difference between two consecutive frames aiming to reduce the computation complexity; while (Kameda & Minoh, 1996 and Collins, et al., 2000) separately proposed a double frame differencing technique for human motion estimation. In (Kameda & Minoh, 1996), the second frame is subtracted from the first frame and the third frame from the second frame, while in (Collins, et al., 2000) the first frame and second frame are subtracted from the third frame respectively. These two techniques are better than the single frame differencing scheme in terms of efficiency, but they cannot completely solve the problem of false positives (i.e. the number of background pixels which are wrongly marked as foreground pixels) and false negatives (the number of foreground pixels which are wrongly marked as background pixels) and may not extract all parts of the object of interest.

In (Zhou & Hoang, 2005) a modified frame differencing technique proposed using a running average for background subtraction, by calculating the average of a number of consecutive frames which would be compared pixel-by-pixel with the current frame using a threshold for similarity to construct a background reference. To solve the problem of false positives in background subtraction, in this technique, the background reference and threshold are updated frequently

To benefit from the advantages of the techniques proposed in (Zhou & Hoang, 2005) and (Mayo & Tapamo, 2009), in (Migliore, et al., 2006) joint differencing is proposed to detect moving objects in video surveillance. To solve the problem of efficiency, this technique employs a hybrid technique, which uses both frame differencing and background subtraction. This technique is not expensive on the computational level, but failed to solve the problem of false positives in some cases such as those caused by waving tree leaves and illumination change. To overcome this problem, in (Heikkilä & Silvén., 2004) a temporal average used to detect the foreground, by calculating the difference between the current frame and background reference, and then the background updates, depending on whether the pixel passes the threshold or not. This technique is also not sensitive to changes in illumination. To overcome this problem, in (Yi & Liangzhong, 2010) a technique proposed to combine the running average and frame differencing. This technique updates the background pixel dynamically, depending on the result of the difference between the current frame and previous frame. Then, if the pixel has not passed the threshold, the background update does not change. By calculating the difference between the current frame and background reference, and taking the frame difference between the current frame and new frame, a moving object can be detected. This technique performs better than the traditional running average, in terms of detecting moving objects.

In (Stauffer & Grimson, 1999) a Mixture of Gaussian model presented, whereby each pixel in the frame represents a mixture of Gaussians (MoG) and updates continuously. This technique aim to model the values of a particular pixel as a mixture of Gaussians. Background colors determined based on the persistence and the variance of each of the Gaussians of the mixture. Pixel values that do not fit the background distributions are considered foreground until there is a Gaussian that includes them with sufficient, consistent evidence supporting it. This technique can provide a good result in terms of efficiency and fixes the problem of gradual illumination change. In (Sen-Ching & Kamath, 2004) a survey conducted for some background subtraction algorithms, and concludes that MoG can designate the object (foreground) more precisely as compared to previous techniques, but it needs more time for processing.

In (Gyaourova, et al., 2003), a block matching algorithm proposed for detecting motion between the current frame and a new frame. The sum of the distance for the grey value in the two blocks is calculated, and then the smallest total distance is

considered as a match. This technique shows the importance of selecting the right block size: a bigger block is less sensitive to noise, while the smaller one produces better results. In addition, a bigger block can solve the problem of aperture, but requires more processing time.

In (Töreyn, et al., 2005) MoG in spatial domain compared with the approximation wavelet subband, and show that the wavelet scheme provide almost the same result as using a spatial domain but with reduced time for processing. The possibility of using more than one level of wavelet for background subtraction is presented in (Aghaee, et al., 2011; Töreyn, 2005 and Tao, et al., 2009).

In (Töreyn, et al., 2005) a technique proposed which used three levels of discrete wavelet transform for all frames, by taking the difference between all coefficients in each frame with the estimated background. This technique does not perform the inverse of the wavelet transform; this is because the result shows that the wavelet domain can provide the same information as the spatial domain. In (Aghaee, et al., 2011) background subtraction technique proposed, by using wavelet to segment objects. In this technique, three levels of WT were used for the current frame and background frame. This method used approximation coefficients (LL) and ignored the remaining subbands, which leads to a reduction in computational complexity and also decreases the noise. In this method, background models are created by calculating the frame average. In (Tao, et al., 2009) a technique proposed using Redundant Discrete Wavelet Transform (RDWT) for detecting the foreground in the scene. In this method, the frame difference is produced by calculating the difference between all coefficient of the subbands in the current frame and the previous frame separately. The sum of the results are then calculated and the pixels, which pass the threshold, are foreground pixels, otherwise they marked as background pixels. In this thesis we shall investigate both spatial domain and frequency domain schemes to eliminate the background. We shall develop those schemes using knowledge gained from existing approaches but incorporating some modifications to deal with efficiency of computation and reduction of positive false errors. These schemes together with experiments to test performance will be presented in Chapter 3.

2.1.2 Gait Cycle Estimation

Background subtraction methods results in binary images/frames, in which background represented by 0 and the foreground (also referred to as silhouette) is represented by 1. After eliminating the background, a bounding box need to be drawn around the silhouette image in each frame. The difference in the size of silhouette images affects the performance of the recognition methods. To overcome this problem in the proposed methods, the silhouette image is extracted according to the size of the bounding box and then normalized to have the same size. It is also necessary to align the silhouettes horizontally to be in the centre of the bounding boxes of successive frames.

The motion of human gait is repeated in the stable frequency gait cycle. Having removed the background and normalised the size of silhouette boxes, the gait cycle is the final step of pre-processing by first determining the sequence of silhouette frames between two identical events while the person is walking. It is usually measured from one heel strike of the ground to the next time the same heel strikes the ground. It is divided into two phases; the stance phase and the swing phase. The stance phase is the first part of the cycle, which comprises 60% of the cycle and consists of 5 components (initial contact, loading response, mid-stance, terminal stance and pre-swing), see Figure 2.1. The swing phase is the second part of the cycle, which comprises 40% of the cycle and is divided into 3 components (initial, middle and terminal swing) (Katiyar, et al., 2010).

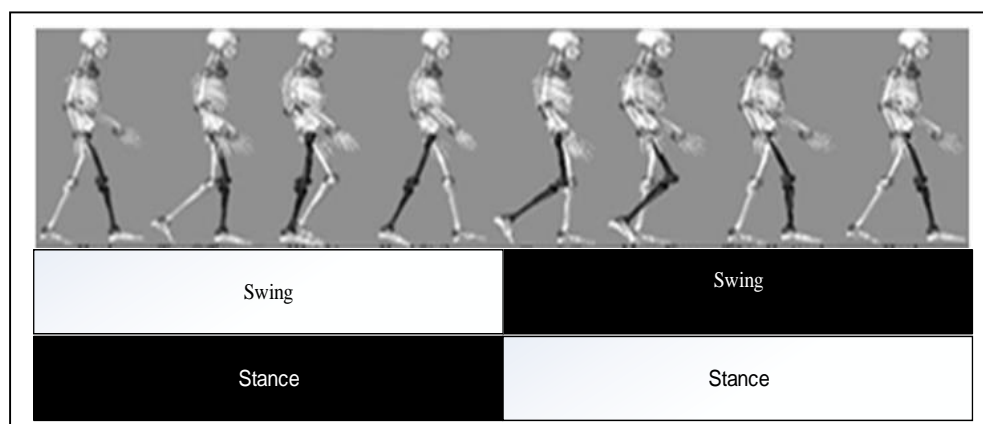


Figure 2-1: The complete gait cycle: stance and swing

In the literature gait cycle were determined mostly based on the local minima and local maxima. Local minima of a width signal correspond to mid-stance positions of a normalized silhouette image sequence. Likewise, the local maxima of a width signal correspond to double support positions of a normalized silhouette image sequence. Thus, the gait cycle length can be determined by calculating a time interval between two consecutive local minima or local maxima of a width signal.

In (Lee, et al., 2010) the gait cycle detected by finding local minima that represent the mid stance of the human silhouettes. In (Sinha, et al., 2013) human gait recognition proposed based on Kinect and a half gait cycle is detected automatically between two consecutive local minima. In (Wang, et al., 2003), estimating the gait cycle is performed using the aspect ratio of the silhouettes bounding box, while in (Sudha & Bhavani, 2011) width of the bounding box is used as it is periodic and the bounding box will be larger and shorter when the legs are farthest apart and thinner and longer when the legs are together. In (Hassin, et al., 2014) gait cycle estimated based on three local minima which generated by computing the aspect ratio of bounding box to all binary image result of pervious step. In (Dikovski, et al., 2014) gait cycle estimated based on three dimensional space between the points of the ankle joints of the subject, a full gait cycle is considered the period between three consecutive local maxima. In (Lee, et al., 2011) foreground sum method is employed to determine the gait cycle period, the process of foreground sum can be permit by first, calculating a number of pixels within the half bottom of a silhouette. Then, the signal generated by a sequence of silhouettes is low-pass filtered. Finally, the consecutive peaks or valleys in the smoothed signal are measured to determine the gait cycle period. In (Bharti & Gupta, 2011) gait cycle estimated based on 4 frames which represent one gait cycle as shown in Figure 2.2. In this thesis, we shall use the local maxima of the bounding box width sequence correspond to double support positions (in which the body weight is supported by both legs). Hence, the gait cycle estimated by calculating a time interval between two consecutive local maxima to determine the gait cycle.

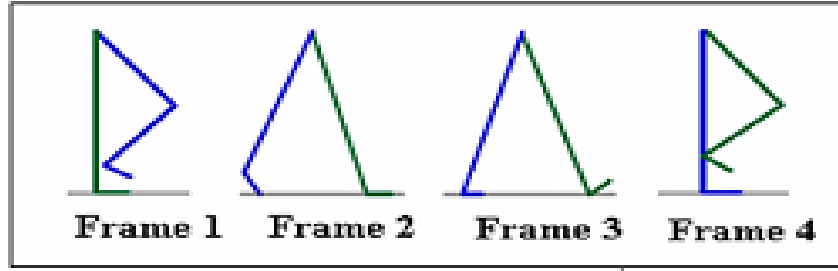


Figure 2-2: Gait Cycle estimated based on 4 frames which represent one gait cycle. (Bharti & Gupta, 2011)

2.2 Gait Feature extraction and Classification

Having removed background and extracted the human gait cycle, gait recognition schemes are characterised by their feature extraction and the classification method. Feature extraction is the process by which key numerical values are determined from an input gait used as the digital feature vector representation of the way the person is walking. It is a crucial step in analysing gait of a walking person any gait recognition scheme, and it is expected to produce vectors in vector space on which a similarity/distance function can be defined for comparison. Following that, a classification method is adopted to enable making identification/verification decisions depending on a criteria, mainly determined through training, to be satisfied by the output from the similarity function.

Feature extraction methods for human gait recognition are classified either as model-based or as motion-based. In this thesis, we use the silhouette which extracted from videos provided by a conventional camera (the CASIA B database), and we investigate the model-based method by extracting features from a skeleton provided by a Kinect sensor. In the next two sections we review the literature on these feature extraction and classification methods.

2.3 Human gait recognition using conventional camera

Ways of recognising individuals by their gait have been widely discussed in the literature for many years. Research by Kumar et al. (Bharatkumar, et al., 1994) led to the hypothesis that gait is a characteristic that can be used as a biometric to identify people. In the early 1970's, medical studies tried to treat gait as a discriminating characteristic, and initial work in this area was carried out by psychologists in 1971, when Johansson attached light points to people's joints in a darkened area. The

candidates were then asked to walk, run or ride a bicycle (Kozlowski & Cutting, 1977 and Cutting & Kozlowski, 1977). Only in the 1990's, that gait recognition was dealt with by machine vision researchers (Venkat & De Wilde, 2011).

Gait features can be captured from a distance, without the knowledge/collaboration of the candidate. In other biometric systems, the candidate needs to participate in the biometric feature capturing process. This factor has motivated researchers to focus on the use of gait recognition in security surveillance systems. Gait recognition can generally be classified into two parts: model-based and motion-based (model-free). Model-based approaches use a structural model to represent human gait (Bharti & Gupta, 2011; Kim, et al., 2009 and BenAbdelkader, et al., 2002). A structural model uses the topology or the shape of human body parts (e.g. head, feet, torso, hip, thigh, knee and ankle). Measuring lengths and distances between different parts of the body on the move has been used as attributes of gait feature vectors. The model-based approach provides us with the ability to directly extract the gait feature vectors (i.e. signature) from the motion of the human body. In (BenAbdelkader, et al., 2002), structural parameters were presented, based on stride and cadence. The cadence is optioned by the walking periodicity, while the stride length is computed by calculating the ratio of the distance travelled and the steps taken. In (Singh & Jain, 2010) dynamic features proposed for representing the human body movements, by constructing two triangles between three points; the first triangle was based on the hands and the right and left heel and the second triangle used the same hand with the right and left toes. Based on these two triangles, a third triangle was constructed by defining the intersecting points of the two triangles. They then calculated the mean value of the angles formed by the new triangle in one gait cycle (see Figure 2.3-(a)). The authors of (Jhapate & Singh, 2011) also relied on dynamic features to represent the human body motion by using one of the hands and both feet as a feature. The concept involves inserting dots on selected body location of the subject in all frames during one gait cycle. A triangle can then be constructed from these points, and by calculating each angle of the triangle, the mean value of one cycle can be determined (see Figure 2.3-(b)). In (Bharti & Gupta, 2011), the extracted features are constructed from four selected frames of the gait cycle, two representing the opening phase of the legs and two representing the closing phase. From each frame, four body points are selected (i.e. palm, knee, ankle and toe) and as a result sixteen nodes (points) are used to

represent the gait cycle. In (Bobick & Johnson, 2001), four different features were used to propose model-based gait recognition. These features were extracted from the calculations of human bodies to create a feature vector. The first feature is obtained by calculating the height of the human body (d_1), the second feature is the distance between the head and the pelvis (d_2), the third is obtained by calculating the distance between the foot and the pelvis (d_3), and the last feature is the distance between the feet (d_4) (see Figure 2.4). Some model-based gait recognition approaches, like in (Wang, et al., 2004), use lower body dynamics as the discriminative features and almost completely ignore upper-body dynamics, due to the difficulty in the accurate extraction of upper-body dynamics. Despite this, in (Lu, et al., 2008) upper-body dynamics (the arms, shoulders and head) were used and they show that these features can provide valuable information for human identification.

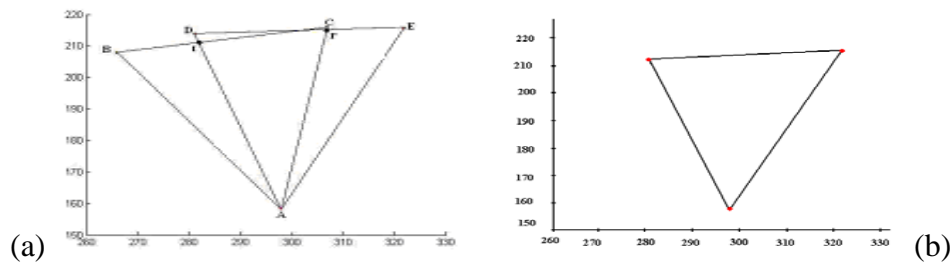


Figure 2-3: Gait signature by creating a triangles from hands and feet (Singh & Jain, 2010)

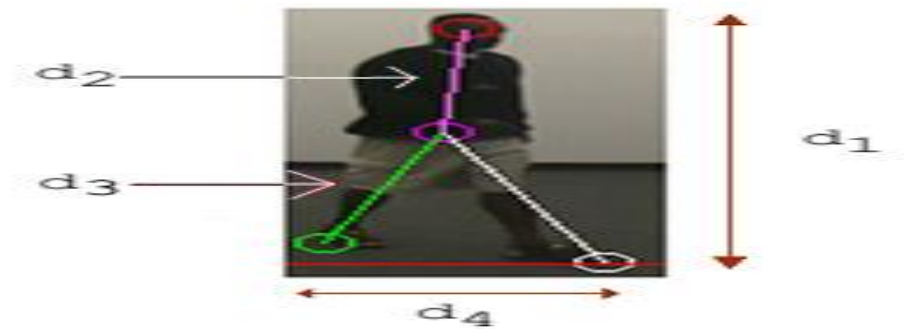


Figure 2-4: Gait signature using static, activity-specific parameters. (Bobick & Johnson, 2001)

In (Wagg & Nixon, 2004), head displacement as a static body parameter is used in addition to the use of lower body parts for recognition purposes. (Wang & Liu, 2007), used a frequency domain gait recognition approach based on body joint position, with this approach, the joint coordinates were calculated using the geometrical appearances

perceived during one gait cycle. Joint coordinates were used to create limb angles and then the Discrete Fourier Transform was applied. With this approach, the feature vector extracted is based on the amplitude frequency and phase frequency of the angles. The weaknesses of the model-based approach have been suggested to include the difficulties in model construction, model fitting and parameter extraction, beside high computational complexity (Preis, et al., 2012), moreover these kind of feature are more difficult in low resolution images especially in real time system due to feature extraction process. On the other hand, motion-based recognition methods have less computational complexity, because they deal directly with the motion pattern of the body, without extracting its principal structure (Lee, et al., 2011; Lee, et al., 2010; Hong, et al., 2007 and Mansour, 2012). Two classes of Motion-based methods are identified in the literature; state-space methods and spatio-temporal methods (Hofmann & Rigoll, 2012). State-space methods deal with the sequence of static body poses (Mansour, 2012), while spatio-temporal methods describe both the spatial and temporal distribution of the human gait (Sudha & Bhavani, 2011). Considering the characteristics of the model and motion based features discussed above, extracting model parameter as a gait signature from silhouettes may help to provide informative features for human gait.

In Motion-based methods, the gait features are extracted from the silhouette. In (Hong, et al., 2007), mass vector is used as a feature, defined as the number of pixels in the rows of silhouettes (i.e. pixels of binary value 1) The coordinates of the mass vectors are dealt with as time-series and accordingly the Dynamic Time Warping (DTW) techniques is used for matching, to overcome the problem of direct frame-by-frame matching. This is due to the fact that in reality people may slightly alter their speed and style of walk. Another motion-based method was proposed by (Lee, et al., 2011) which is based on step-forward and step-following templates. In this technique, static body shape and dynamic motion were used to represent the Frame Difference History Image (FDHI). In (Arantes & Gonzaga, 2011), Global Body Motion features are proposed by using four different features; the Silhouette-Grey-Wavelet model (SGW) representing grey motion; the Silhouette-Binary-Wavelet model (SBW), representing binary information, the Silhouette-Edge-Binary model (SEW), encapsulating edge information; and the Silhouette Skeleton Wavelet model (SSW) representing motion in terms of skeletal movement. In (Foster, et al., 2001), area-based

metric system used for measuring gait signatures. Gait masks are placed over the silhouette, each corresponding to a specific part of the human body image, and the change in the selected area of the masks are measured and used as a feature vector.

A Gait Energy Image (GEI) is a model-free gait representations that is frequently proposed as a gait feature. GEI can be compute based on the average silhouette images to represent both body shape and movement during one gait cycle. GEI is also known to be an efficient gait feature and overcomes the problem of low quality silhouette images. In (Han & Bhanu, 2006), GEI is proposed, whereby synthetic templates and statistical methods were used for gait recognition purposes. Conditional-Sorting Local Binary Patterns are proposed as a method in (Hsia, et al., 2013), this method uses local binary patterns to extract gait features based on GEI. GEI has been investigated in neutral gait sequences and it has been shown to achieve good performances. However, in the case of non-neutral conditions, like carrying a bag or wearing a coat, these methods are very sensitive. Some method attempts to remove bags and coats on human silhouettes, like in (Pratheepan, et al., 2009), which proposed dynamic and static feature templates based on GEI, this method outperforms methods which use traditional GEI. Further research has tried to improve GEI to reduce the effect of these covariate factors. In (Sivapalan, et al., 2012), Backfilled Gait Energy Image (BGEI) proposed as another gait energy based feature that can be generates from both; side view silhouettes and frontal depth images, this method allows the feature to be applied across different capturing systems using the same enrolled database. In (Li & Chen, 2013), foot energy images and head energy images were combined to generate a new technique called structural gait energy image (SGEI). Similar to GEI, silhouette edge information is used to create one image during one gait cycle. (Hofmann & Rigoll, 2012), used a method called Gradient Histogram Energy Image (GHEI), the gait features extracted based on the calculation of gradient histograms at all locations of the original image. In (Xu, et al., 2012) GEI represented as a set of local augmented Gabor features, which concatenate the Gabor features extracted from different scales and different orientations together with the X–Y coordinates. In (Hsia, et al., 2013), the Conditional-Sorting Local Binary Pattern (CS-LBP) technique was proposed, based on GEI, in order to focus on the critical information of human gait. A Gait Entropy Image (GEnI) is another silhouette based future which was recently used to represent gait signature, GEnI feature aim to generate an image based on the

randomness of pixel values in the silhouette images over a complete gait cycle (Bashir, et al., 2009). Due to the high dimensionality that is produced by GEI techniques, some techniques use dimension reduction methods to reduce the dimensions of extracted feature vectors before classification. In (Chourasiya, et al., 2013 and Ali, et al., 2011), Principal Components Analysis (PCA), as a dimension reduction technique applied on GEI. Furthermore, in (Liu & Sarkar, 2006) Linear Discriminant Analysis is used with the Hidden Markov Model to represent gait features based on shape space. Some existing work on gait recognition has dealt with one variation only. For example, in (Martín-Félez, et al., 2011) the neutral gait recognition method is proposed, based on different types of GEI, by segmenting the gait cycle into four biomechanical poses (Double limb Stance, Initial Single Stance, Mid-Stance and Terminal Single Stance). By doing so, this generates a particular GEI for each pose. In this method, the CASIA B gait database was used to test its performance, and the Leave-One-Out method was used to estimate the performance of each strategy, followed by a nearest neighbour classifier. In (Yu, et al., 2014), gait recognition is proposed based on three types of optical flow (Lucas-Kanade, Horn-Schunk and Multi-frame). In this method, the CASIA B gait database was used, and the experiment was conducted based on neutral and non-neutral conditions separately. For the neutral gait sequence, the Leave-One-Out method, followed by nearest neighbour, was used to test its performance.

In the real world, there is no guarantee of having only one variation in the gallery and probe sets. We may face different variations, and for this reason, examples in literature show other techniques that attempt to provide an overview of different variations in their proposed methods. In (Li & Chen, 2013), gallery sets include neutral gait sequences only, and probe sets contain different variations, like carrying a bag and wearing a coat, while in (Bashir, et al., 2010), as addition this scenario another scenario also presented by having different gait sequences variations in the gallery. In (Hu, et al., 2013), gait recognition is proposed under non-neutral conditions. The gait signature was extracted based on the LBP flow (LF). In this method, LBP is used to describe the texture information of optical flow and the dynamics of the gait signature are represented by the Hidden Markov Model (HMM), which is beneficial for both tracking and recognition. The recognition process in this method is divided into two parts; the first part is model-based, which tries to generate a statistical model for all candidates, compared with the probe set, this method compares individual HMM

(iHMM) with the classical offline HMM (o HMM). The second part is called exemplar-based, in this method the distance between gallery samples and probe data directly performed and the average different sequence (AVG) and dynamic time warping (DTW) are used to construct the feature. In (Jeevan, et al., 2013), the Gait Pal and Pal Entropy (GPPE) method was used to present the sensitivity of such a method to changes in various covariate conditions, such as carrying an item and clothing, this method used PCA as a dimension reduction followed by Support Vector Machine (SVM) as a classification method. This method was compared with Shannon entropy (ShGEnI) to highlight its efficiency. In (Bashir, et al., 2010), the feature selection based on Gait Entropy Image (FGENI) is proposed for human gait recognition, which ignores the effect of covariate factors aiming to overcome the problem of covariate conditions. To reduce the high dimensions of the extracted feature vector this method used Adaptive Component and Discriminant Analysis (ACDA). In (Yogarajah, et al., 2011), a Poisson Random Walk of Gait Energy Image (PRWGEI) is proposed, which aims to overcome covariate factor effects on gait features as well, and the feature vector reduced by Linear Discriminant Analysis (LDA) aiming to improve the discriminative power of the extracted feature vectors. The space-time interest points method (STIPs) is another method which proposed in (Kusakunniran, 2014) to deal with gait recognition under covariate gait sequences, based on both spatial and temporal direction. This method is based on motion information rather than global shape information, which aim to be more robust in dealing with covariate factors (wearing a coat or carrying a bag). In this method first The Space-Time Interest Points (STIPs) are detected from a raw gait video sequence and then the Histogram of Image Gradient (HOG) and Histogram of Optical Flow (HOF) used to provide information from each point in the STIPs.

Considering the proposed gait feature in the literature, to reduce the effect of covariate factors (such as, coat and bag), we expect that the use of Detail wavelet coefficients subbands (LH, HL and HH), in which the edge of the human body is provided may facilitate the recognition of gait. Moreover some of the mentioned proposed features represent different characteristics of human gait. Fusing some of those features at different levels of fusion might improve the performance of gait recognition system. In this thesis we aim to investigate different kind of feature vectors and fusing them at different levels.

2.4 Human gait recognition using Kinect sensor

Recently, researchers have been focusing on using the Kinect sensor, which can detect the joint points of a human body more precisely. Kinect is a gaming device, manufactured by Microsoft for the X-box-360. Its primary function is to respond to horizontal human interaction and movement, utilising a sensor that is used in different applications, such as games, healthcare and security surveillance. The authors of (Parajuli, et al., 2012) used a Kinect sensor to present a health monitoring system to collect data on posture recognition (sitting Vs standing) and for gait recognition (normal walking Vs abnormal walking). In (Preis, et al., 2012), biometric features were extracted to represent the human body in motion, using 12 static features (i.e. height, torso, both lower legs, right and left upper arms, the length of both legs, both thighs and both forearms), and 2 dynamic features (i.e. the step length and speed). They asked 9 persons to walk from right to left side view of the Kinect sensor, and from each candidate 8 trails were recorded. The 14 extracted features were classified into three different sets of features; the first set included the height, length of legs, length of the torso and length of the left upper arm; the second is static features; and the third set used only dynamic features. Three different classifiers were used to test the performance of this method called: R1, C 4.5 and Naive Bayes. The first is aim to generates a classification rule based upon a single feature in the training data, the second is run based on a decision tree, while the third is depend on the probabilistic classifier based on the Bayes' law. The first set of features with R1, C 4.5 and Naive Bayes achieved 62.7%, 76.1% and 90 % rates of recognition respectively, which comprises the best set of features compared with the recognition rates of the other two sets. The authors of (Ball, et al., 2012) proposed a larger list of 18 dynamic features, based on the changing angle in the lower limb joints, and calculated the mean, maximum and standard deviation of the three angles for both legs in the half gait cycle. The data set used for unsupervised clustering in this method contained four people which provides a total of 71 samples of walk half-cycles. K-means with a Euclidean distance metric were used to classify the candidate and this method obtained 43.6% as recognition rate. Other researchers (Sinha, et al., 2013) proposed 2 sets of features detected by the Kinect sensor; the first set included static features comprising joints from the upper body (i.e. the centre of the shoulders, the left shoulder, the left hip, the centre of the hips, the right hip and the right shoulder) and from the lower body (i.e.

the centre of the hips, the right hip, the right knee, the right ankle, the left ankle, the left knee and the left hip). The second set consists of 4 distances, measured between the hands and legs. Taken as the dynamic features, the Adaptive Neural Network technique was used for selection and classification. The experimental result based on this method shows that the static features get higher importance compared to the dynamic features. A proposal using 2.5D view invariant gait recognition was presented in (Tang, et al., 2014) based on point cloud registration using the Kinect sensor and the feature vector created based on Gaussian curvature and mean curvature. This method highlight the problem of view direction by proposing the multiple-view gait recognition. In short, most of the literature on the use of Kinect camera exploit the variety of features that are detected readily by the sensor and argue for a specific set/sets of features. In this thesis we will investigate two sets of dynamic feature, to propose two methods based on neutral and non-neutral (coat wearing and carrying bags) gait sequences using Kinect sensor.

2.5 Gender classification based on gait features

In general, gait features can be analysed and used for two main purposes; gait recognition, where a subject's ID is identified in a specific environment and gait classification, which includes gender classification, action classification and estimation of age (Chen, et al., 2009). There is a lot of research dealing with human gait recognition, but only a few recent works have used gait features for gender classification. Gender classification can be applied for different purposes, e.g.: (1) surveillance systems ; (2) gender classification may be also used as a first step in some biometric recognition like gait to improve the accuracy, or when searching gait databases based on one gender only, etc. In (Mather & Murdoch, 1994), body sway, waist-hip ratio and shoulder-hip ratio are used as features for gender classification, because males tend to swing their shoulders more than their hips, and females do the reverse. The human silhouette model is used by different researchers in different ways for gender classification purposes. The Embedded Hidden Markov Model (EHMM) is proposed in (Chang & Wu, 2010) for gender recognition using a 2-Dimensional Discrete Cosine Transform (2D-DCT) coefficients as gait feature vectors. In (Lee & Grimson, 2002 and Martín-Félez, 2010 'A') the gait feature were generated from moving human silhouettes for gender classification. Each silhouette in a sequence is

divided into seven regions and an ellipse is fit to a region. The seven areas were corresponds to: head/shoulder region; front of torso; back of torso; front thigh; back thigh; front calf/foot; and back calf/foot, as shown in Figure 2.5.

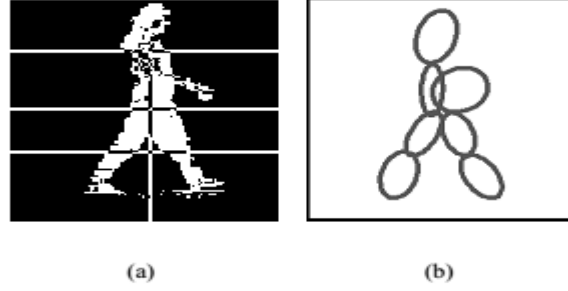


Figure 2-5: Dividing the silhouette into 7 regions, and ellipses example to extract gait feature to be used for gender classification (Lee & Grimson, 2002).

While in and (Martin-Felez, et al., 2010 'B') the body silhouette is divided into eight more realistic regions: head, torso, two arms, two thighs and two calves/feet. In this method the torso is fully segmented, and if the two arms are visible, then they independently located.

In (Yu, et al., 2009), the Gait Energy Image (GEI) is divided into five regions: head/hairstyle, chest, back, waist/buttocks and legs, to be used as feature vectors. Different influence factors, such as clothing and carrying condition changes, were addressed in (Hu, et al., 2010) by proposing a hierarchy approach for gait-based gender classification. This technique used low dimensional discriminative representation, obtained as the Gabor-MMI feature. In addition, gender related Gaussian Mixture Model-Hidden Markov Models (GMM-HMMs) were constructed in this classification model. In (Arai & Asmara, 2011), the energy of 2D Discrete Wavelet Transform is used for presenting gender recognition, based on gait feature. The method proposed based on, HL (Vertical), LH (Horizontal) and HH (Diagonal) subbands were used to extract gait energy during one gait cycle, in this method 6 level decomposition energy used to extract a feature vector. All the detail coefficients were combined to get one row of a new data, then a combination of Horizontal and Vertical coefficients performed. The experimental result in this method shows that classification accuracy is 90.5 % using combination of Detail coefficient, while is 92.9% using the combination of Horizontal and Vertical coefficient. A relative way of GEI is proposed

by (Arai & Asmara, 2012), called Gait Energy Motion (GEM). In this method the motion feature presented as spatial data from human gait. (T. Anitha, February 2013) Proposed another gait-based gender classification and the gait signature extracted by dividing the silhouette image into two regions (upper and lower). The first region related to upper part starting from head to the torso region whereas the second region is from torso to the feet. Ellipse is formed for each region and the parameters were measured to represent the feature vector. In this method all the 31 females and 31 randomly selected males are used in the experiments. To test the performance of their method they used kNN and SVM as a classification method.

In the CASIA B database the number of samples of male and female is not balanced (93 males and 31 females). Most of the reported gender classification studies select 31 Male samples randomly ones and ignore the effect of others, to be used with the 31 female samples for gender classification model as in (Yu, et al., 2009 ; T. Anitha, February 2013 and Arai & Asmara, 2012). However, this might not presenting a reasonable measure for the model performance, because the result may change while using another random set of 31 male samples. In (Martin-Félez, et al., 2010 'A'), and (Martin-Félez, et al., 2010 'B') 31 male samples selected 10 times randomly to give more chance for the whole 93 male samples to participate in the probe and gallery set. In this thesis gender classification method applied on CASIA B database. To provide more reasonable approach, we use 25 male and 25 female samples randomly and this process is repeated 30 times, and finally the mean and standard deviation computed and presented.

2.6 Chapter Summary

In this chapter we presented description of the background for gait recognition supported by a survey of existing research investigations and proposals that deals with the various components of gait biometrics including background removal, gait cycle detection, feature extraction and ending with classification approaches. We are now in a position to present our investigations and development of gait recognition scheme(s) during the rest of the thesis. In the next chapter we shall present our investigations that focuses on the pre-processing and feature extraction.

Chapter Three

Human gait Recognition System

The process of gait recognition in this thesis has been developed by proposing various methods, to present these method we divide this chapter into four main phases. The first phase is pre-processing, explained in section 3.1 which cover background subtraction (i.e. object extraction) and gait cycle estimation. For Background subtraction we investigate different background subtraction methods, and present the results of experiments to test the performance of each method. This would be followed by gait cycle estimation and normalisation. In section 3.2 we investigate feature extraction in terms of different feature vectors that proposed for gait recognition including; Spatio-temporal Model (STM), Legs Motion Detection (LMD), Approximation coefficients Wavelet Model (AWM) and Detail coefficients Wavelet Model (DWM), Vertical Distance Feature (VDF) and Horizontal Distance Feature (HDF). We further investigated the use of different sets of features extracted from the commonly used Gait Energy (GEI) and Gait Entropy (GEnI) compactification images of the gait sequences captured from different part of the body.

The rather high dimensional nature of the extracted feature vectors is dealt with, in this thesis, by dimension reduction and feature selection. In section 3.3 we present three methods; Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) and Fisher score. The last part of those chapter cover the different classification methods (k-nearest-neighbour (k-NN), Support Vector Machine (SVM) and Linear Discriminant Classifier (LDC)) and describe the used databases.

3.1 Pre-processing

Pre-processing is an important step in the process of gait recognition, the aim of pre-processing is to improve the image (frame) data which overwhelms undesired distortions or enhances some image features related to further processing and analysis task. In our proposed gait recognition methods the pre-processing step includes: background subtraction, normalisation and gait cycle estimation

3.1.1 Background Subtraction technique

Background subtraction is a main step of pre-processing in gait recognition, in chapter-2 we presented different methods that proposed to deal with background subtraction. In this section we will investigate four methods of background subtraction and test their performances on a selection of videos. The first two methods are existing methods and the difference between them is that the first has difficulties in dealing with outdoor video recordings. We slightly modify second method but the modified version still suffers from inefficiency. We developed the third and fourth methods to improve efficiency over the 2nd method. In the first method we used four video recorded by ourselves in indoor scene while the other method were taken from Action database.

3.1.1.1 Action Database

This database was created for human actions, including walking and running. To create this database, various actions were performed at different times. It included 25 subjects in four different scenarios, outdoors and indoors, and different clothes were also worn. The created database was based on six types of human actions (walking, jogging, running, boxing, hand waving and hand clapping). All sequences were captured with a homogeneous background using a static camera with a 25fps frame rate (Schuldt, et al., 2004). In this thesis, the Action database was used for the background subtraction method.

3.1.1.2 Background Subtraction – Motion Compensation

This method is based on the spatial domain, using frame differencing and background references. In this method the process of background subtraction is begins by converting the RGB frames to grey. Then, to create a background reference we calculate the average for the first ten frames, assuming that there is no image in the scene, then by calculating the difference between consecutive frames with the background reference, we produce the foreground image and draw a box boundary around the foreground, (see Figure 3.1).

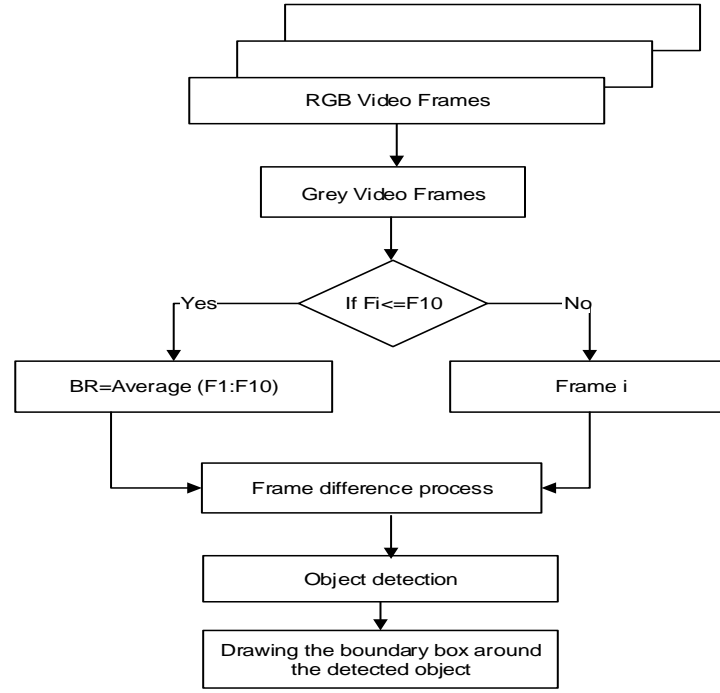


Figure 3-1: Background subtraction Motion Compensation (F is the frame and $i = 1: N$, N is the number of frames)

Results

The above method was tested on four video sequences taken from our own video sequences which recorded in indoor scenes with 320×240 image resolution and 25 frames/second. This technique provide significant result in terms of time needed for processing and it is more suitable for databases were recorded in indoor scenes as compared to outdoor. This method can provide (0.13) second /frame as a time that needed for processing (see Table 3.1).

Table 3-1: Background subtraction- Motion Compensation process time using four video sequences taken from our own indoor video sequences.

Video No.	Second/frame
1	0.1439
2	0.1137
3	0.1386
4	0.1239
Average	0.13002

The drawback of this method is that it has difficulties in dealing with outdoor video sequence. In the next method we present MoG that deal with indoor and outdoor video sequences.

3.1.1.3 Background Subtraction – MoG

To further improve the efficiency of background subtraction, in this method we try to reduce false positive (the number of background pixels which are wrongly marked as a foreground) and false negative (the number of foreground pixels which are wrongly marked as a background). The proposed method is based on MoG which has been proposed before frequently. MoG can extract the object from background efficiently, but it has the drawback of computation complexity. To fix such a problem, we divided the frame into five parts, then we only used the first part for processing, until the object come to the scene. Then, the frame width (column) will increased step by step, after which we drew a box boundary around the object (human) (see Figure 3.2).

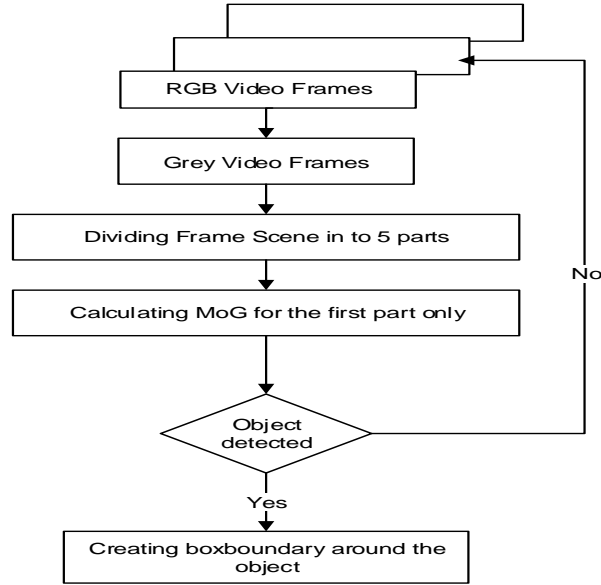


Figure 3-2: Background Subtraction based on MoG. First each frame divided vertically into five parts, and then MoG applied on first part to reduce the computation complexity.

Results

We implement this experimental based on four different video sequences taken from action database. This method was implemented based on direction only. For this experimental 320×240 frame resolution was used with 25 frames/second. This method evaluated only in terms of processing time, because we exploit MoG for being a widely used method and the modification is made using a part of frame rather than the whole frame to reduce the computation complexity. The experimental result shows that this method can isolated the foreground without needing to apply MoG on the

whole frame size which reduce the time that needed for processing, this method can provide (0.65273) second/frames as an average of processing time (see Table 3.2).

Table 3-2: Background Subtraction –MoG, process time based on action database

Video No.	Second/frame
1	0.6548
2	0.6265
3	0.64514
4	0.6845
Average	0.65273

These results show that price of improved performance for outdoor, much richer background, recording is the significantly increased background removal time. This increase will make timely recognising people from their gait on the move an unrealistic option. The next method will attempt to improve efficiency.

3.1.1.4 Background Subtraction –MoG in the wavelet domain

Researchers have recently focused on the use of wavelet transform in background subtraction, this is to reduce the computational complexity and it also provides very robust results under illumination changes. Numerous of background subtraction techniques focus on using approximation coefficients in wavelet transform. Researchers claimed that the difference in detailed coefficients images can provide more stable detection results that is why in this method we proposed a technique based on a detailed coefficients for background subtraction. To decide whether the scene contains the object or not, we used approximation coefficients. The main steps of this method are as follows:

Step 1: After converting the frames from RGB to grey, we applied wavelet transform for grey frames. The first frame is used as a frame reference, then we calculate the difference between the first frame and consecutive frames. For this step, we used the approximation coefficient only.

Step 2: We applied standard deviation (SD) for the frame reference (FR) and the differences between the frame reference and consecutive frames (DFR). If the standard deviation of the DFR frame is greater than the standard deviation of the FR, we carried on with the process, otherwise we do not calculate step 3.

Step 3: In this step, we created a matrix based on the size of one coefficient subband, then we checked each pixel in the three non-approximation coefficients (LH, LV and HH). If the pixel in all non-approximation coefficients passed the threshold, then we used the maximum value in the new matrix, otherwise we made the pixel zero in the new matrix. To create the threshold, we used normal distribution, because the non-approximation coefficient is represented as Laplacian distribution.

Step 4: In this step, we used the mixture of Gaussian method as the background subtraction technique.

Step 5: In final step, we use obtain the contour of the object (silhouette). (see Figure 3.3).

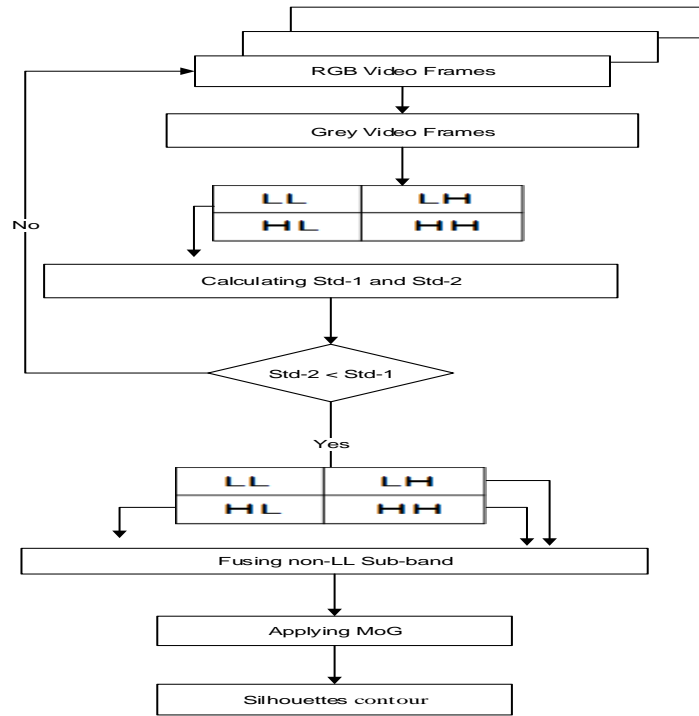


Figure 3-3: Background subtraction in the wavelet domain, using non-approximation coefficients (LH, LV and HH). Std-1 is the standard deviation of the frame reference and Std-2 is the standard deviation of the difference between consecutive frames with frame reference.

Results

This method implemented based on four video sequences taken from action database and the experimental results shows that the proposed method can provide significant result in term of efficiency and computation complexity as well. In table 3.3 we present the process time of this method using four different video sequences with 320×240

frame resolutions with 25frames/second. This method needed (0.13475) second / frame as a time for processing one frame (see Table 3.3). Normally human body represented vertically based on this fact we aim to improve the efficiency of this method by using the Vertical coefficient instead of using the combination of the three non- approximation coefficients (LH, LV and HH).

Table 3-3: MoG in the wavelet domain process time, using the combination of the three non- approximation coefficients (LH, LV and HH), based on action database.

Video No.	Second/frame
1	0.137
2	0.137
3	0.130
4	0.135
Average	0.13475

3.1.1.5 Background Subtraction – Refined MoG in the wavelet domain

In this method we proposed a technique based on the vertical coefficient subband only, by using the following steps:

Step 1: After converting the frames from RGB to grey, we applied wavelet transform for grey frames. The first frame is used as a frame reference, then we calculate the difference between the first frame and consecutive frames. For this step, we used the approximation coefficient only.

Step 2: We applied standard deviation (SD) for the frame reference (FR) and the differences between the frame reference and consecutive frames (DFR). If the standard deviation of the DFR frame is greater than the standard deviation of the FR, we carried on with the process, otherwise we do not calculate step 3.

Step 3: In this step, we created a matrix based on the size of vertical coefficient subband, then we checked each pixel in the vertical coefficients. If the pixel is passed the threshold, then we used the value in the new matrix, otherwise we made the pixel zero in the new matrix. To create the threshold, we used normal distribution, because the non-approximation coefficient is represented as Laplacian distribution.

Step 4: In this step, we used the mixture of Gaussian method as the background subtraction technique.

Step 5: In final step, we use obtain the contour of the object (silhouette).

Results

In this method we aimed to propose this method to provide better result in term of efficiency and computation complexity as well, because human represented vertically in this method we used only Vertical coefficient, by this we can get most necessary information of human couture. For applying this technique we used four video sequences from the action database, in this method 25 frame/second used with 320×240 frame resolution. This method can provide (0.123325) second / frame as average time for processing (see Table 3.4).

Table 3-4: Refined MoG in the wavelet domain process time, using Vertical coefficient, based on action database.

Video No.	Second/frame
1	0.1259
2	0.1224
3	0.1257
4	0.1173
Average	0.123325

In this thesis we used action Database to test the background subtraction methods only. CASIA B gait database and Kinect database 1 and 2 were used to test the performance of gait recognition method. In Kinect database we don't need to use background subtraction because the object is already detected. Although CASIA B database, has provided silhouettes, but unfortunately some of them are missing and some are not extracted properly. For this reason we used the original videos provided by the database to enhance the silhouette extraction process and produce new set of silhouettes.

Although, the modified MoG procedures is successful in outdoor as well as indoor. In the rest of the thesis we only used Motion Compensation for background removal due to the fact that current work is restricted to indoor gait recognition. However, the inclusion of the new wavelet-based MoG algorithm was originally for efficiency comparison purposes but it is motivated by our future plans to extend the current schemes to outdoor situation.

3.1.2 Gait Cycle estimation and Normalisation

The gait cycle is the term describing the ambulatory phase of walking which starts from a certain heel-strike of a single foot to the same heel-strike again. The gait cycle is divided into two parts; stance phase and swing phase. The stance phase is the first part of gait cycle which is 60% of the cycle and consists of four components (loading response, mid-stance, terminal stance, and pre-swing). To extract the gait cycle from human gait various ways were proposed as mentioned in chapter-2. Because the width of boundary box changes to be larger or shorter, when the legs are widely separated or overlapped, In this thesis we determine the gait cycle using local maxima based on the width of boundary box, by finding double support position of the silhouettes, the boundary box approach is shown in Figure 3.4. To make the process of feature extraction easier, we normalize the silhouette images to the same size. All the silhouettes have been resized to have the same height and the normalized silhouettes will be aligned to the horizontal centre.

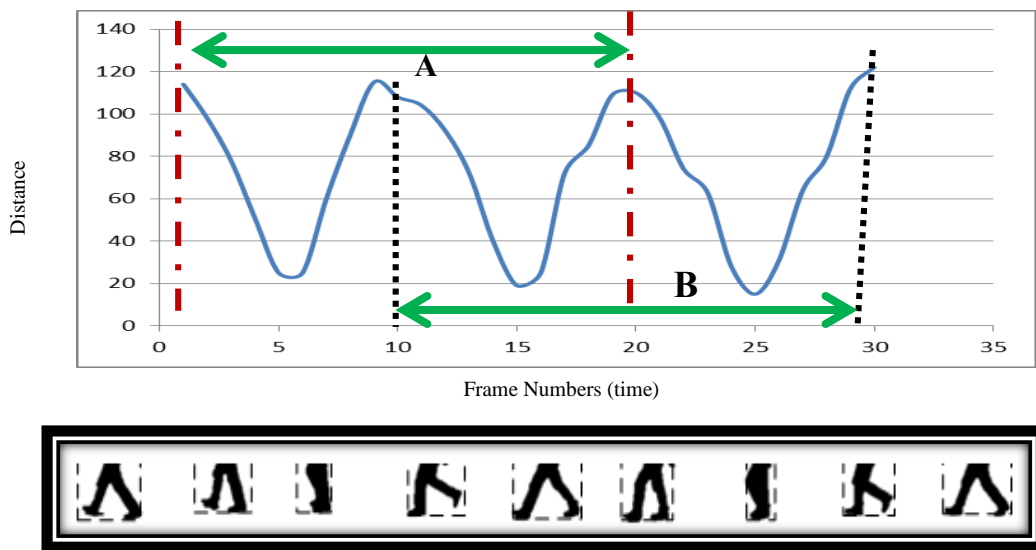


Figure 3-4: Gait cycle estimation overview

3.2 Feature Extraction

Feature extraction is the process by which key features of the samples are selected or enhanced. Typically, the process of feature extraction is a crucial step in analysing gait as a part of any method of biometric recognition. The methods used in feature extraction are different depending on the type of biometric recognition. Feature extraction in human gait recognition can be done using two methods; a model-based

method and a motion-based method (see chapter-2). Feature extraction in this thesis is presented based on two different parts; the first part focuses on features, which are extracted from videos provided by a conventional camera (CASIA B database in proposed thesis), and the second part deals with the model-based method, which extracts features from a skeleton provided by a Kinect sensor. We present each of these type separately. From the CASIA B six main feature vectors will be extracted called STM, LMD, AWM, DWM, GEI and GEnI. To extract these feature vectors, we first isolate the foreground from the background using motion compensation, then the gait cycle will be estimated as mentioned earlier. For the first three set of features STM, LMD, and AWM, we do the following steps; firstly we extract human body from the grey frame by recalling pixel values that have been detected as a foreground (mapping). Secondly we draw the boundary box around the detected foreground then we apply three level of wavelet transform. We extract STM and LMD from LL1, while from LL3 we extract AWM and finally from Level 1, 2 and 3 of non-LL wavelet subbands we extract DWM. GEI and GEnI will be extracted from spatial domain and then wavelet transform applied on the constructed GEI and GEnI image. In the next subsection we will describe each of the 5 feature vectors mentioned above.

3.2.1 Spatio-temporal Model

In this method, a set of specific body movements is extracted to represent special features. The distances between hands ($d1$), shoulders ($d2$), feet ($d3$), and the height of the person ($d4$) were extracted from the human body during a gait cycle, based on the LL1 subband (see Figure 3.5).

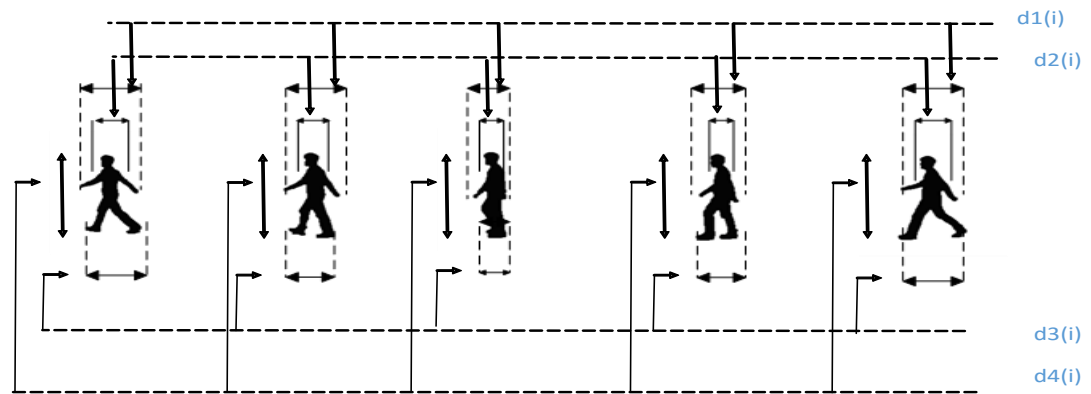


Figure 3-5: Spatio-temporal Model (STM) feature vector, using the distances between hands ($d1$), shoulders ($d2$), feet ($d3$), and the height of the person ($d4$).

To detect each of the above distances we used golden ratio proposition (Herman, 2007), followed by boundary box, to remove the unwanted part from human body. To detect the distance of the feet, the upper body part that represent 0.62 of the person's height will be removed, while to detect the hands distance, we remove 0.38 of the person's height from lower body part. Finally to detect the distance between the shoulders we remove 0.67 of person's height from lower body part. After that we draw the boundary box around the remaining part of human body to detect the distance accurately, (See figure 3.6).

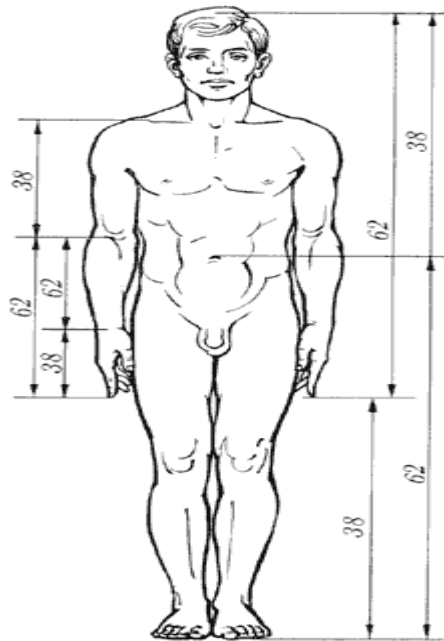


Figure 3-6: Golden ratio proportions for human body.

<http://www.redbubble.com/people/coldblood/journal/5228103-composition-golden-ratio>

The combinational spatial features of human body silhouettes for consecutive frames in one gait cycle can be used to represent a temporal feature. In this method, we extract the spatial features, based on the distances presented in Figure 3.5. These distances are extracted from each frame during one gait cycle. Min, Max, Mean, skewness and Standard Deviation were calculated for each of d1, d2, d3, and d4 separately during one gait cycle. Minimum and maximum distances represents two different models of the human body; firstly, when the legs are completely overlapping, as shown in Figure 3.7 (a), and secondly, when the legs are wide apart,

as shown in Figure 3.7 (b). The first model is depending on the minimum distance of the hands, shoulders and feet with the maximum height. Conversely, the second model is represented by the maximum distance of the hands, shoulders and feet with the minimum height. Moreover, calculating mean, skewness and standard deviation represents the temporal changes of distance in one gait cycle. The way the feature vectors are constructed is presented mathematically as follows:

$$MinD = \{\min(d1_{i=1:n}), \min(d2_{i=1:n}), \min(d3_{i=1:n}), \min(d4_{i=1:n})\} \quad (3-1)$$

$$MaxD = \{\max(d1_{i=1:n}), \max(d2_{i=1:n}), \max(d3_{i=1:n}), \max(d4_{i=1:n})\} \quad (3-2)$$

$$MeanD = \{\text{mean}(d1_{i=1:n}), \text{mean}(d2_{i=1:n}), \text{mean}(d3_{i=1:n}), \text{mean}(d4_{i=1:n})\} \quad (3-3)$$

$$StdD = \{\text{std}(d1_{i=1:n}), \text{std}(d2_{i=1:n}), \text{std}(d3_{i=1:n}), \text{std}(d4_{i=1:n})\} \quad (3-4)$$

$$SkewD = \{\text{skew}(d1_{i=1:n}), \text{skew}(d2_{i=1:n}), \text{skew}(d3_{i=1:n}), \text{skew}(d4_{i=1:n})\} \quad (3-5)$$

$$STM = \{MinD, MaxD, MeanD, StdD, SkewD\} \quad (3-6)$$

Where d is the distance, i is the frame number and n is the number of frames in one gait cycle and STM is the 20-dimensions of spatio-temporal feature vector used in gait recognition.

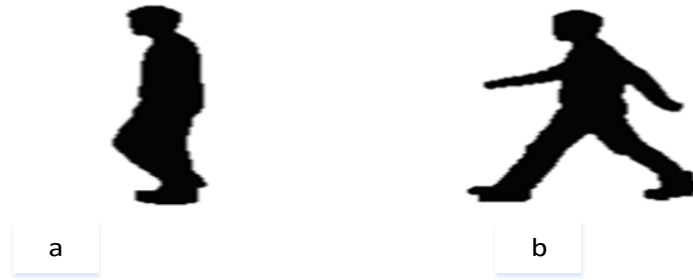


Figure 3-7: (a) Human body when the legs are overlapping, (b) human body when the legs are wide apart.

3.2.2 Legs Motion Detection

To present the motion of the human legs during one gait cycle we propose another feature vector called LMD. This feature was extracted from each leg separately (see Figure 3.8). Firstly, the legs are detected based on the human body proportions given by the golden ratio (Herman, 2007), using $(\text{height} \times 0.38)$ to be used for all frames (N)

during one gait cycle. The normalized legs area were divided vertically into two parts (Rear and Frontal), to detect the pixel that change in each part, the hamming distance is calculated and saved for any consecutive frames during one gait cycle. For the rear side leg, the feature vector is called the Rear Leg Hamming Distance (RLHD), while the feature vector based on the frontal leg called the Frontal Leg Hamming Distance (FLHD). The mean is calculated, in order to detect the motion changes during one gait cycle for RLHD and FLHD separately. In order to find the variance of motion change, standard deviation (Std) is applied for each side. Finally, in order to detect the change between consecutive frames during one gait cycle, we applied the method of least squares (ls). LMD mathematically defined as follows:

$$\text{MeanH} = \{\text{mean}(\text{RPCC}_{i=1:n}), \text{mean}(\text{FPCC}_{i=1:n})\} \quad (3-7)$$

$$\text{StdH} = \{\text{std}(\text{RPCC}_{i=1:n}), \text{std}(\text{FPCC}_{i=1:n})\} \quad (3-8)$$

$$\text{LsH} = \{\text{ls}(\text{RPCC}_{i=1:n}), \text{ls}(\text{FPCC}_{i=1:n})\} \quad (3-9)$$

$$\text{LMD} = \{\text{MeanH}, \text{StdH}, \text{LsH}\} \quad (3-10)$$

Where RPCC and FPCC are the Rear and Frontal Legs Pixel Change Counter respectively, i is the frame number, n is the number of frames in one gait cycle and LMD is the Legs Motion Detection of 6-dimensions feature vector.

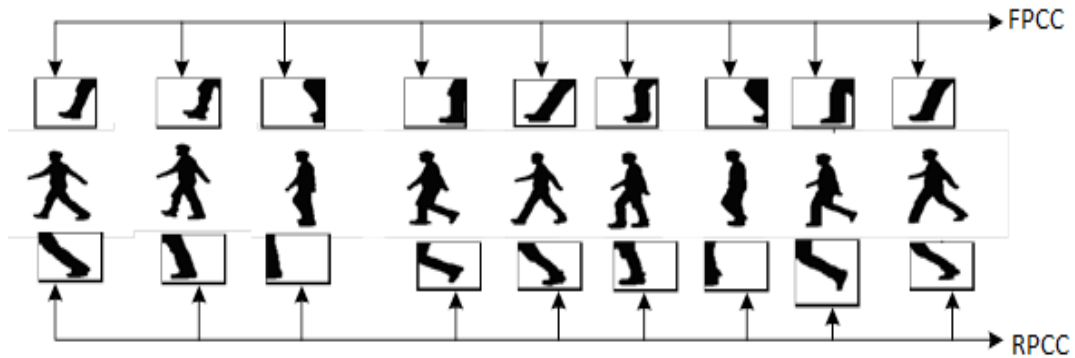


Figure 3-8: The overview of Legs Motion Detection (LMD) feature vector using Rear and Frontal Legs Pixel Change Counter

3.2.3 Approximation and Detail coefficients Wavelet Model

In this model, we applied Discrete Wavelet Transform (DWT) using the Haar wavelet. First we mapped the boundary box around the grey colour of the human body in the

frames (detected based on background subtraction) and then wavelet transform applied on the detected part. In order to extract this feature the human body will be divided into two parts – the Upper Body Part (UBP) and the Lower Body Part (LBP), based on the human body proportions provided by the golden ratio (see Figure 3.9). ($UBP = 0.62 \times Height$) and ($LBP = 0.38 \times Height$). Two different feature vectors are constructed in this model; the first is extracted from the LL subband, while the second is constructed from Detail coefficients (LH, HL and HH). In the LL subband statistical property (mean and standard deviation) of each frame during one gait cycle were calculated and entitled the mean set and standard deviation set (in Detail coefficients only the standard deviation calculated for each frame during the gait cycle). After constructing the two previous sets, the mean and standard deviation were calculated for each set, in order to construct the feature vector. When the process is based on the LL-subband, the constructed (8-Dimensional) feature vector is called AWM conversely, when the process is based on the detail coefficients subband (LH, HL and HH), DWM with 36-Dimensions is constructed as a feature vector (more detail in chapter 5). The process of AWM is calculated as shown below:

$$mLBP^i = \text{mean}(LBP_i) \quad (3-11)$$

$$mUBP^i = \text{mean}(UBP_i) \quad (3-12)$$

$$sLBP^i = \text{std}(LBP_i) \quad (3-13)$$

$$sUBP^i = \text{std}(UBP_i) \quad (3-14)$$

$$MLBP = \text{mean}(mLBP_{i=1:n}) \quad (3-15)$$

$$MUBP = \text{mean}(mUBP_{i=1:n}) \quad (3-16)$$

$$SLBP = \text{std}(sLBP_{i=1:n}) \quad (3-17)$$

$$SUBP = \text{std}(sUBP_{i=1:n}) \quad (3-18)$$

$$MSLBP = \text{mean}(sLBP_{i=1:n}) \quad (3-19)$$

$$MSUBP = \text{mean}(sUBP_{i=1:n}) \quad (3-20)$$

$$SMLBP = \text{std}(mLBP_{i=1:n}) \quad (3-21)$$

$$SMUBP = \text{std}(mUBP_{i=1:n}) \quad (3-22)$$

$$AWM = \{MLBP, MUBP, SLBP, SUBP, MSLBP, MSUBP, SMLBP, SMUBP\} \quad (3-23)$$

Where UBP and LBP are the Upper and Lower Body Parts respectively, n is the number of frames in one gait cycle.

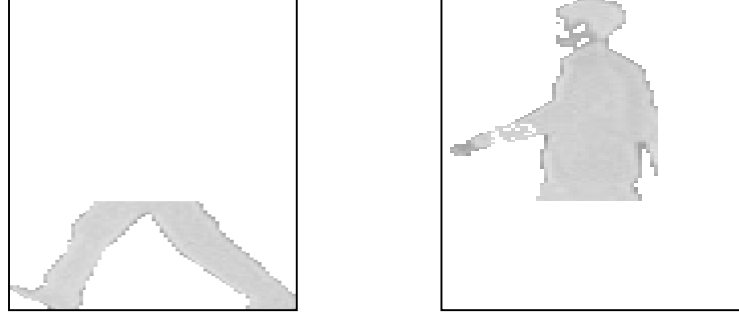


Figure 3-9: AWM example (a) Lower human body part

(b) Upper human body

3.2.4 Gait Energy Image

GEI is the time-normalised build up energy image of human gait in one gait cycle. GEI tries to obtain average silhouette images to represent both body shape and movement over one gait cycle (see Figure 3.10).

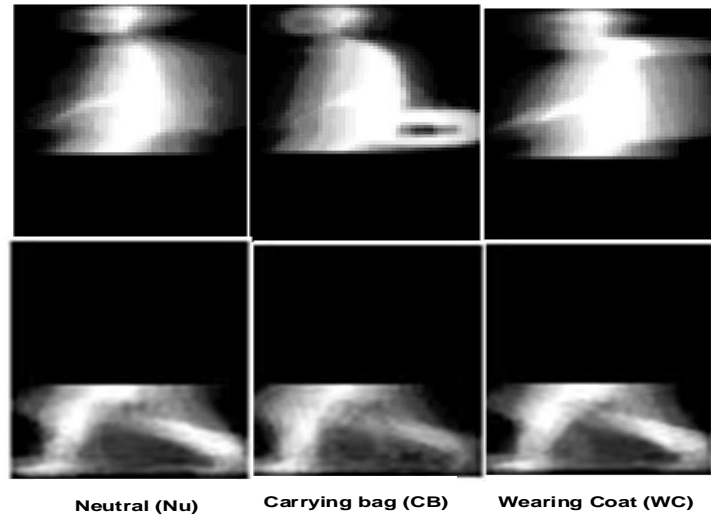


Figure 3-10: Examples of Gait Energy Image, for Neutral gait (Nu), Carrying Bag (CB) and Coat Wearing (CW), from CASIA B database.

GEI has been used in different research related to gait recognition and it has been shown to be an effective gait feature; it is easy to compute and is insensitive to noise

in silhouette extraction (Han & Bhanu, 2006). Pixels with the highest intensity values in GEI represent a lack of noticeable motion in this part during a gait cycle, while pixels with less intensity values represent the gait motion of the human body. GEI, like most existing gait representations, deals with the motion and appearance information. The drawback of GEI is its sensitivity when dealing with large intra-class variations (e.g. wearing a coat and carrying conditions). GEI has been proposed by different studies in different ways, as mentioned in the literature review. In this thesis, we propose the use of GEI in another way, to deal with gait recognition under covariate factors (Nu, CB and CW), based on various sets of features.

$$G(x, y) = \frac{1}{T} \sum_{t=1}^T B(x, y, t) \quad (3-24)$$

Where T is the number of frames in a complete cycles of the sequence, t is the frame number in the sequence, x and y are values in the 2D image coordinate. B(x,y,t) represents the pixel (x,y) in frame t.

In general we have extract 3 feature vector based on GEI called; Lower GEI (LGEI); Upper GEI (UGEI); Whole GEI (WGEI). The upper and lower body are determined during one gait cycle by the golden ratio proportion (Herman, 2007) ($UBP = 0.62 \times Height$) and ($LBP = 0.38 \times Height$), see Figure 3.6.

$$GEI^{WBP} = G(WBP) \quad (3-25)$$

$$GEI^{UBP} = G(UBP) \quad (3-26)$$

$$GEI^{LBP} = G(LBP) \quad (3-27)$$

$$GEI_{feature} = \{ GEI^{WBP}, GEI^{UBP}, GEI^{LBP} \} \quad (3-28)$$

Where WBP, UBP and LBP are the whole, upper and lower body part respectively.

3.2.5 Gait Entropy Image

GENI is a single image constructed from the randomness of pixel values in normalised silhouettes over a complete gait cycle. Entropy aims to ignore the redundancy in the silhouette frames and represent the silhouette of one gait cycle in a single image (see Figure 3.11).

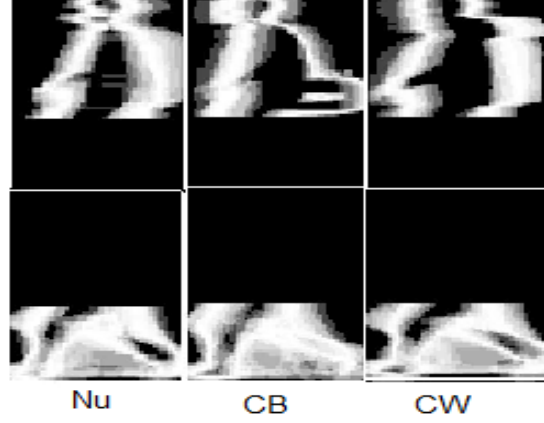


Figure 3-11: Examples of Gait Entropy Image, for Neutral gait (Nu), Carrying Bag (CB) and Coat Wearing (CW), from CASIA B database.

In this method, Shannon entropy is applied to three different models of the human body separately; the whole human body, the upper body and the lower human body. The proposed GEnI is mathematically defined as follows:

$$E(x, y) = - \sum_{r=1}^R p_r(x, y) \log_2 p_r(x, y) \quad (3-29)$$

Where x and y are the pixel coordinates, $p_r(x, y)$ is the probability of the value r (1 or 0) in the sequence of the (x, y) coordinate value along the whole frames. In the case of having binary silhouettes images, like in our case, so we have $R = 2$ (Bashir, et al., 2010). A $GEnI(x, y)$ can then be obtained by scaling $E(x, y)$ so that its value ranges from 0 to 255 as follows.

$$GE(x, y) = \frac{(E(x, y) - E_{min}) * 255}{(E_{max} - E_{min})} \quad (3-30)$$

In GEnI we will also divided the human body into three parts Lower (LGenI); Upper GEnI (UGEnI); and Whole GEnI (WGenI), using the same way proposed in GEI.

$$GEnI^{WBP} = GE(WBP) \quad (3-31)$$

$$GEnI^{UBP} = GE(UBP) \quad (3-32)$$

$$GEnI^{LBP} = GE(LBP) \quad (3-33)$$

$$GEnI = \{ GEnI^{WBP}, GEnI^{UBP}, GEnI^{LBP} \} \quad (3-34)$$

In chapter 6 we shall describe other features that can be obtained directly from the Kinect sensor. These include Horizontal distance features and vertical distance features associated with gait cycle and extracted by the camera using 20 automatically detected body skeleton points (see Figure 3.12).

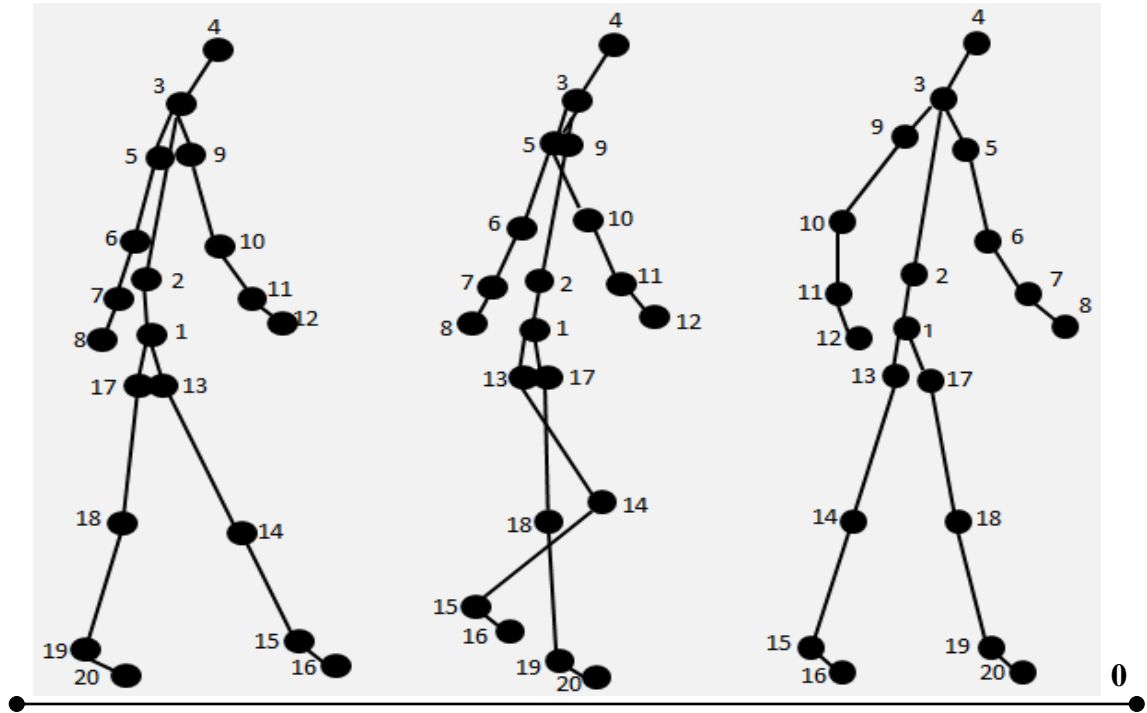


Figure 3-12: Overview of the skeleton points provided by Kinect Sensor.

3.3 Dimension reduction

The various feature extraction components, described above, results in a high dimensional representation of the gait biometric trait. High dimensionality of the feature space in gait features introduces computational and statistical challenges. The nature of high-dimensional data in any pattern recognition tasks requires the establishment of dimension reduction or dimension selection procedures that enables us to efficiently analyse the data adequately without lose of discriminating power. Reducing the dimensions of the data usually means selecting a basis for a relatively low dimensional subspace of the original high dimensional space, within which you can describe most of the variances within your data. The main objective of dimension reduction techniques is to remove irrelevant and redundant data, thus reducing the computational complexity and avoiding data over-fitting. The process of reducing the attributes subset of a feature set can be divided into Feature Subset Selection and

Feature Transformation techniques (Addison, et al., 2003). Feature Selection is the systematic manual selection of relevant feature attributes, from which the original characteristics are maintained; while Feature Transformation is the generation of linear combinations of the original characteristic set. In this thesis, both Feature Selection and Feature Transformation were proposed to address the aforementioned challenges.

3.3.1 Feature Transformation

The process of transforming high dimensional data into lower dimensional feature space is known as Feature Transformation. The feature transformation may be a linear or non- linear combination of the main feature. Feature Transformation approaches are decisive methods in image pre-processing and analysis of hyperspectral data. These methods allow the system to be more efficient on a computational level (Zamalloa, 2008).

In this thesis, we used two common methods for Feature Transformation; PCA and LDA, to reduce the dimensions of the feature vectors. Both of these feature transformation methods (PCA and LDA) aim to extract features by projecting the original parameter vectors into a new feature space, out of a linear transformation matrix using different transformation methods.

3.3.1.1 Principal Component Analysis

PCA is a statistical procedure used for transforming high dimensional data into low dimensional data, while retaining most of the original information. The main advantage of PCA in biometric systems is to reduce the dimensionality of the features, by finding a new set of variables, with fewer dimensions than the original set of variables (Jolliffe, 2002). The variables generated after the process of transformation are called Principal Components (PCs). In chapters 4 and 5 of this thesis, PCA will be used for the purpose of dimensionality reduction in gait recognition. Mathematically, PCA can be presented as follows:

Assume that we have an n-dimensional feature vector called “a” for data $\{a_i | i=1, 2 \dots N\}$, which can be stated as a matrix $A = [a_1, a_2, \dots, a_N]$. The aim is to summarise them by projecting into r-dimensional subspace.

Firstly, the mean value is calculated for every column of K as follows. If $A_{n \times d}$ is a matrix of data with n samples and d feature. The Covariance matrix can be defined as:

$$C_{d \times d} = A^T A \quad (3-35)$$

The direction with highest variance can be found by computing the Eigenvalues and Eigenvectors of the matrix C .

$$Cx = \lambda x \quad (3-36)$$

A set of the Eigenvalues $[\lambda_1, \lambda_2, \dots, \lambda_d]$ is used to find the corresponding Eigen vectors each in d dimension, which will represent the PC projection $d \times d$ matrix, where their columns X_i are sorted decreasingly based on their correspondence Eigenvalues.

3.3.1.2 Linear Discriminant Analysis

LDA is a method that seeks to reduce dimensionality by generating a linear combination of variables that gives the best possible separation between groups (classes) (Duda & Hart, 1973). It has been used widely in many applications, including high-dimensional data, such as face recognition and gait recognition. LDA aims to reduce high-dimensional feature vectors belonging to different classes to a lower dimensional feature space. Thus, the projected feature vectors of a class on this lower dimensional space are well separated from the feature vectors of other classes. LDA attempts to calculate the sample mean of each class. Then, it calculates the sample covariance, by first subtracting the sample mean of each class from the observations of that class, and then by taking the experimental covariance matrix of the result.

The main requirement of LDA is that the sample size must be greater than the length of the feature vector. In this thesis, LDA will be used in chapters 4 and 5 as a dimension reduction method.

Let m_i , $i=1,2,\dots,c$ is the mean of the sample belong to class i and c is the number of class and let m be the mean vector of d -dimension for the whole data samples, where d is the number of feature.

Here we shall define the within classes scatter matrix as:

$$S_W = \sum_{i=1}^c S_i \quad (3-37)$$

Where

$$S_i = \sum_{j=1}^{n_i} (x_j - m_i) T (x_j - m_i) \quad (3-38)$$

And

$$m_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_j \quad (3-39)$$

Where n_i is the number of sample in c_i Between class

$$S_B = \sum_{i=1}^c (m_i - m) T (m_i - m). \quad (3-40)$$

The final scatter matrix need to be maximized is represented as:

$$A = S_B S_W^{-1} \quad (3-41)$$

3.3.2 Feature Selection

Feature selection is usually used as an alternative to dimension reduction either when the dimension is not very high or when there are different types of features and there are some obvious dependents among some attributes. Feature Selection (FS) is the process of selecting a subset of given features to be used in a model structure. FS methods can be classified into “wrapper” methods and “filter” methods. Wrapper model techniques evaluate features using a learning algorithm, which will ultimately be used to “wrap” the selection process around the learning algorithm. Filter-based approaches always rely on the class labels, most commonly assessing correlations between features and the class label. Although the dimensions of the feature vector used in this thesis are not very high, Fisher score as a feature selection method used to reduce the dimension of the fused feature vectors and to select the most relevant feature we use the. A Fisher score deals with the maximum score of the feature, whereby the distances amongst data points in different classes are as large as possible, whilst the distances amongst data points in the same class are as small as possible (Duda & Hart, 1973). Given a set of data points with label, $\{x_i, y_i\}_{i=1}^n$, $y_i \in \{1, \dots, c\}$ where c is the number of classes. Let n_i denote the number of data points in class i . Let μ_i and σ_i^2 be the mean and variance of class i , corresponding to the r -th feature. Let μ and σ^2 denote the mean and variance of the whole data set. The Fisher score is defined as below:

$$FS(f_i) = \frac{\sum_{j=1}^c n_j (\mu_{i,j} - \mu_i)^2}{\sum_{j=1}^c n_j \sigma_{i,j}^2} \quad (3-42)$$

Note that the Fisher score is based on the same theoretical model of the dimension reduction LDA scheme.

3.4 Classification Methods

Classification methods aim at assigning a class label, from within the relevant classes, for each input example. The performance of object classification schemes in general rely on the nature of the deployed image sensors; the adopted image pre-processing schemes; and the reliability of the object detection, object segmentation, and the feature extraction model (Kamavisdar, et al., 2007)

To classify persons we need a database that contains predefined patterns that compares with detected person to classify in to proper category. There are many methods proposed for classifying person in gait recognition as any other biometrics. This section briefly describes the various classification methods used in the thesis in order identifying\classifying persons based on their gait.

3.4.1 k-Nearest-Neighbour

k-NN is one of the most essential and simple classification methods. It is simple to realise and works very well in practice. It is more useful compared to other classification methods when there is little or no prior knowledge about the distribution of the data. Another advantage of k-NN classification is that it is not expensive on a computational level. k-NN methods needs distance metrics method to calculate the distance between two input vectors which can be define mathematically as following:

$$x_i \text{ and } x_j, x_i = (x_{i1}, x_{i2}) \text{ and } x_j = (x_{j1}, x_{j2}). \quad (3-43)$$

$$(x_i, x_j) = |x_i - x_j| = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2} \quad (3-44)$$

The distance between two k-dimensional vectors, $x_i = (x_{i1}, x_{i2} \dots x_{ik})$ and $x_j = (x_{j1}, x_{j2} \dots x_{jk})$ can be calculated as follows:

$$D(x_i, x_j) = |x_i - x_j| = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 \dots + (x_{ik} - x_{jk})^2} \quad (3-45)$$

The k-NN classifier attempts to provide a class membership as an output, based on the majority vote of its neighbours. If $k = 1$, then the object is simply assigned to the class of that single nearest neighbour. Nearest neighbour classifiers are similarity-based classifiers; the outcome is based on comparison; both test object data and class labels are matched and compared to similar training data. Finding a nearest neighbour is dependent on distance functions, and the latter's judgement is conducted by assessing the similarities and differences between data objects (Tomasev, et al., 2011). Further the classification accuracy of KNN is also basically dependent on the distance metrics methods used for computing the nearest distance. In this thesis in order to compute the optimal distance metric we used: Euclidean distance and City Block.

3.4.2 Support Vector Machine

SVM is a machine learning algorithm, which is widely used as a powerful technique for classification methods. SVM is basically used to deal with data with two classes, known as binary classification. It is also used for multi-class data, but this is more expensive on a computational level (Meyer, 2004). SVMs are able to cope with non-linear solutions efficiently using kernel methods. The process of SVM is based on the concept of decision planes that define decision boundaries. In SVM, a decision plane is used to separate a set of objects having different class memberships. It attempts to maximise the margin between the classes. Support vectors are data points closest to the separating hyper-plane and the margin is the width between the boundary (Burges, 1998) (see Figure 3.13). In this thesis, we will use SVM for gender classification in chapter 4.

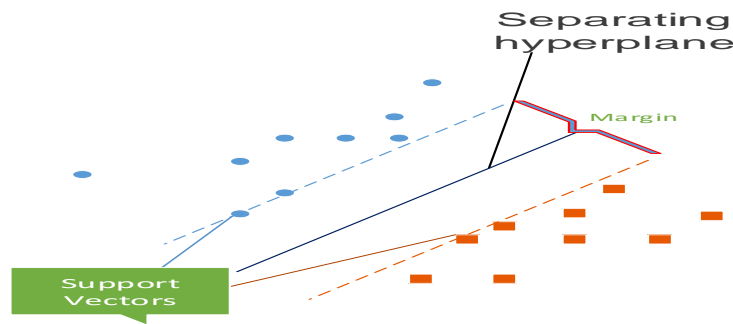


Figure 3-13: Support Vector Machine Overview

3.4.3 Linear discriminant classifier

LDC is a statistical method that aim to find linear combination of the features that best discriminates the data in the classes of interest. These different combinations are called discriminant functions. It is computationally attractive as compared to other classifiers like artificial neural network. LDC belongs to the generative group of classifiers, the classes are expected to have normal distributions and equal covariance matrices (Duda & Hart, 1973). Based on this assumption, the optimal classifier reduces to calculating linear discriminant functions (see Figure 3.14).

The LDC classifier assumes that the population π_i of the class C_i , where $i = 1, 2, \dots, K$ of each class follows a multivariate Gaussian distribution:

$$f(x|\pi_i) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} e^{\left[-\frac{1}{2}(x-\mu_i)'\Sigma^{-1}(x-\mu_i)\right]} \quad (3-46)$$

Where μ_i is the mean vector of the class C_i and samples are classified to class π_i with largest value of $p_i f(x|\pi_i)$ where the monotonicity of the log function implies that $\log(p_i f(x|\pi_i))$ is equivalent to $p_i f(x|\pi_i)$. Here, p_i is the prior probability of the class i . The linear score function is re-expressed as follows:

$$S_i = -\frac{1}{2} \mu' \Sigma^{-1} \mu + \mu' \Sigma^{-1} x + \log p_i \quad (3-47)$$

$$S_i = a_{i0} + \sum_{j=1}^K a_{ij} x_j + \log p_i \quad (3-48)$$

Where:

$$a_{i0} = -\frac{1}{2} \mu' \Sigma^{-1} \mu \quad (3-49)$$

$$a_{ij} = jth \text{ element of } \mu' \Sigma^{-1} \quad (3-50)$$

The priori probability of population i is estimated based on prior knowledge of the class' distribution. In this study we substitute the population covariance by the pooled covariance C_p of the whole classes' data. Thus the score function representing the classification rule is set as follows:

$$S_i(x) = -\frac{1}{2} \mu' C_p^{-1} \mu + \mu' C_p^{-1} x \quad (3-51)$$

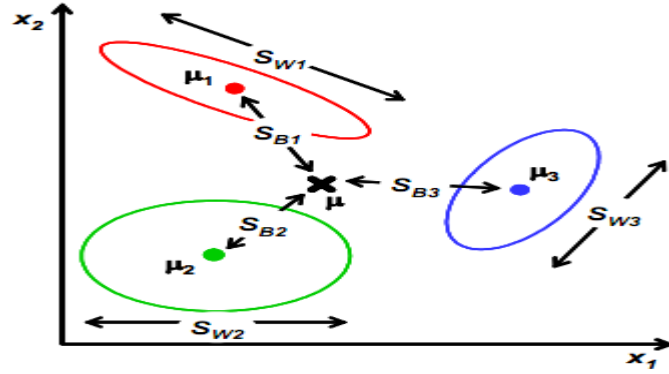


Figure 3-14: Linear Discriminant Classifier Overview. (blog.csdn.net)

3.5 Databases

Testing the performance of any pattern recognition model/scheme, as is the case of the work in this thesis for gait biometrics, requires the use of some benchmark databases that contain sufficiently large and population representative samples of recordings relevant to the investigated task. Performance testing normally takes the form of partitioning the data samples into a gallery and testing sets according to certain established protocols. Databases for gait biometric, must capture the essence of the targeted task in order to evaluate the relevant performance of the system and comparing it with other systems that use the same database. A good database for gait recognition must consist of different variations in the recording condition and environment; the following are databases that take this approach. There are different database for gait recognition such as USF, SOTON and CASIA dataset. In this thesis different databases have been used to test the performance of the proposed methods called Action, CASIA-B and Kinect Database, each of them will be describe in the coming chapters separately.

3.6 Chapter Summary

In this Chapter, we described briefly what has been done throughout the chapter as background preparation, this is including background subtraction, feature extraction, dimension reduction and classification methods. In the next chapter we will study the neutral gait sequences and we will focus on the reliability of the proposed gait recognition decision, finally we will investigate gender classification.

Chapter Four

Neutral gait sequences

Chapter 1 described different challenges when dealing with person recognition using gait that could arise as a result of differences in external factors that may influence personal style of walking gaits during the gait capturing process. We label external factors as non-neutral if the person is carrying bags (on their back, or by hand), or wearing coats. In order to understand the influence of non-neutral factors, we should first deal with the case of neutral conditions which means that the person is not carrying objects or wearing long coats that prevent detection of the full length of legs. This chapter will be devoted to this task by investigating the neutral gait sequences with the aim of gait recognition accuracy rate when using the basic feature vectors extracted from the person's silhouette as described in chapter 3. We shall also discuss the reliability of recognition decision made by the proposed gait recognition decision. We shall also investigate the use of gait information for gender classification, identifying sets of gender relevant gait features. Finally, we design and test the performance of a 2-level neutral gait recognition scheme which begins with gender recognition. We shall demonstrate that using gender information improves recognition accuracy.

4.1 Person Recognition from Neutral Gait Sequences

This section first presents our proposed algorithm for human gait recognition using spatio-temporal body biometric features extracted from the wavelet domain. This will be followed by a description of the experimental dataset and protocols. We finally present and discuss the results of our experiments.

4.1.1 The Neutral Gait Recognition (NGR)

The system has been developed in three main phases (see Figure 4.1). The first phase starts by pre-processing the captured video, which is expected to contain a single individual walking through the scene, having the camera recording at a perpendicular direction to the person's path from one side (the left side in this case). The pre-processing starts with background subtraction is applied using the motion compensation method described in chapter 3, Section 3.1.1. All cropped frames are

then normalized and the gait cycle estimated (see Section 3.1.2). The final step of phase 1 is recalling the grey values that have been detected as foreground. The second phase is the wavelet transforms decomposition of each cropped image resulting from the previous phase, down to level 3. Here we are proposing three new sets of gait features extracted from different wavelet subbands as described earlier in chapter 3 Section 3.2: the first and second feature vectors are extracted from the low frequency subband LL1 called Spatio-Temporal Model (STM) and Legs Motion Detection (LMD), while the third feature vector is extracted from the LL3 and called Approximation coefficients Wavelet Model (AWM). All three feature vectors are then concatenated together for feature-level fusion. Phase 3 consists of the final 2 steps of the algorithm are dimension reduction followed by classification. Linear Discriminate Analysis (LDA), and Principal Component Analysis (PCA) were applied separately to reduce the dimensionality of the feature vector. For classification, we used and compare the performance of the Linear Discriminant Classification (LDC) and the Nearest Neighbour (NN). The following is the block diagram of this 3-phase algorithm.

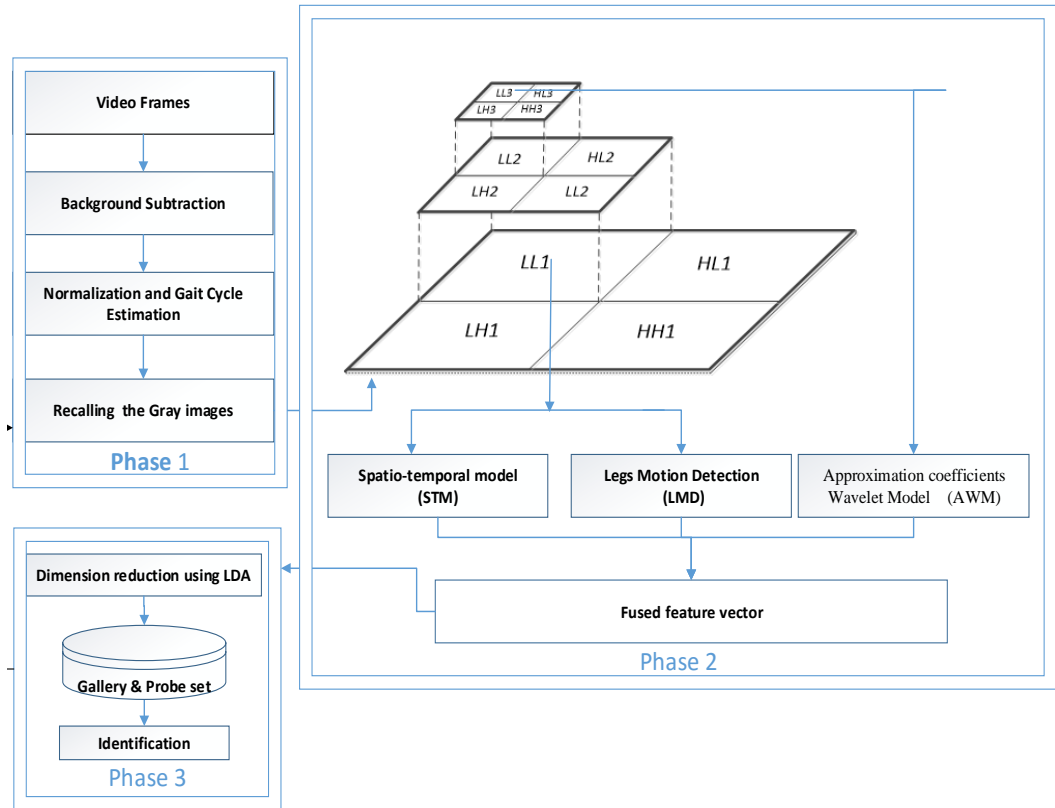


Figure4-1: Overview of the Neutral Gait Recognition (NGR) system.

To test the performance of this NGR scheme we next describe the experimental setup.

4.1.2 The Experimental dataset - CASIA Database

The CASIA gait database is one of the commonly used databases consists of three datasets. The first dataset is called the National Laboratory of Feature Recognition (NLPR) Gait Database, also known as CASIA A. For this dataset, a digital camera was fixed to a tripod and used in an outdoor environment, where a single human moves into view without obstruction. All subjects walk freely in a straight line. Three different viewing angles were used for this dataset, with respect to the image plane; namely frontal (90°), lateral (0°) and indirect (45°). The CASIA A gait database includes 20 different subjects and 4 sequences per view per subject. In total, the database includes 240 ($20 \times 4 \times 3$) sequences. Twenty-five frames per second, with 24-bit full colour images were used and the resolution was 352×240 .

The second dataset is known as CASIA B which is a large multi-view indoor gait database that includes 124 subjects (31 women and 93 men), recorded from 11 different view angles. For each subject, there are 10 walking sequences, consisting of 6 Neutral sequences (Nu), 2 Carrying Bag sequences (CB) and 2 Coat Wearing sequences (CW). The recordings come in three types of variations: varied view angles, wearing different clothing, and variable carrying condition. The indexing labels for the recoded files reflect these variations. In addition to the original video sequences, CASIA B provides sequences of frames from these videos but with the person's silhouettes.

The third dataset is known as Dataset C: this dataset was collected with the use of a thermal imaging camera. It includes 153 subjects and different walking speeds were used; normal walking, slow walking, fast walking and normal walking with a bag. The videos were all recorded at night (Yu, et al., 2006). In this thesis, we shall only use the CASIA B dataset for experimental work on gait recognition and gender classification based on gait.

4.1.3 Experiment Setup

Our current experiments use CASIA B gait database. We only used the side view of the neutral gait sequences in this experiment, in order to provide the richest information with regard to the human gait. In this experiment four neutral walking videos per person were selected as the gallery set, and the remaining two neutral

videos were used as the probe set. In order to provide a realistic performance we calculated the recognition rate based on repeating the experiment so that all the possibilities of selecting 4 neutral gait sequences out of 6, to make up the gallery set. The performance of the gait recognition scheme is determined by the average recognition result of all 15 possible gallery sets. In the whole experiments throughout this thesis we use Matlab2014a 8.4 software for simulation.

4.1.4 Experimental Results

We tested the performance of 7 versions of our system, where in the first 3 schemes a person's gait is represented by the 3 single feature types LMD, STM and AWM. For the other 4 schemes we fused, at the feature level, every two of the three feature types as well as all of them. Figure 4.2 illustrates the recognition accuracy rates for each of the 7 gait schemes. As can be seen the three single schemes LMD, STM and AWM achieved accuracy rates of 29.23%, 68.82 and 50.4% respectively. These rather disappointing results indicate low, but valuable, discriminating powers of each of these feature vectors with the LMD having the lowest discriminating power by a big margin. The tradition in pattern recognition, when one has multiple schemes that are not giving desirable accuracy is to fuse some or all of the schemes. Fusion can be done at different level and we chose the feature level fusion where by the gait is represented by the concatenation of a combination of these feature types. The fused schemes of LMD+STM, LMD+AWM, STM+AWM and all yielded 70.83%, 51.55%, 83.74% and 81.52. These results confirm the benefits expected from fusion except for the last All-features fused system. The best fused system STM+AWM achieves a reasonably good results which is still not anywhere near the state of the art which is in the neutral case is in the upper 90s (see Table 4-2 later). The fact that the STM+AWM scheme outperform the All-features scheme indicates that within the All-features the LMD attributes are redundant and perhaps linearly dependents on combinations of the attributes of the STM+AWM. Note that when we fused the low performing LMD scheme with each one of the two others we obtained a slightly better accuracy on the highest performing feature type indicating that some, but not all, of the LMD attributes are possibly linearly dependent on some of attributes of each of the other types. The low accuracy rate for the LMD, as well as their impact on fusion, can possibly be due to the fact that the LMD features are extracted from the legs only. However, these

results should not be taken to mean that the lower body parts cannot discriminate gaits but rather that the LMD features many not be suitable as they are. We shall come to that question later in Chapter 5.

Note that without involving muscles from arms and upper body parts, leg movements become very slow and it becomes difficult to differentiate gaits of two slow moving persons.

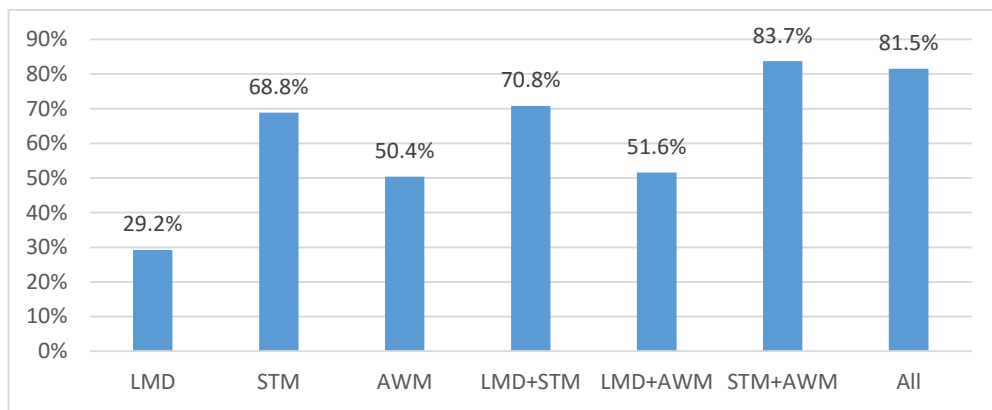


Figure 4-2: The Recognition Performance of NGR using different feature vectors with having 6, 20, and 8 dimensions for each of LMD, STM and AWM respectively. The method tested based on CASIA B gait database (Using 124 subjects with 6 Neutral gait sequence) and k-NN with (k=1) is used as a classification method.

The rather low performance of each of the three types may indicate that for each of these feature vectors there are significant variations in some of the attributes which may have unnecessarily large effects on the Euclidian distances between different samples, while the samples less spreading effects. So for our feature vectors perhaps can test the idea of finding an appropriate linear transformation that may lead to improved performance when doing verification/identification in the transformed domain. In many pattern recognition applications it has been shown that the LDA based schemes outperform PCA based schemes due to the fact that the PCA suffer from inter- and intra-class errors.

We tested our system using LDA transformed version of each of the 7 schemes defined above. Figure 4.3 shows the recognition accuracy rates for schemes included the various fused ones. These results first illustrate the significant benefits we obtained from the use of the LDA transform except for the LMD. The neutral gait recognition

accuracy of the LDA transformed STM and AWM schemes increased to 78.6% and 82.3% respectively, while LMD discriminating power has improved by a very negligible amount of 0.2%. Although LMD provided a lower recognition rate, however when fused with AWM the result increased to 87.5%. It seems that the LDA transform of these two feature vectors assist each other because both are based on body motion. However LMD has a bad effect on STM when fused post LDA transformation, whereby accuracy has reduced to 75.8%. This means that even after the LDA transformation, STM features remained highly correlated to the LMD features. Fusing STM and AWM, on the other hand achieves a high accuracy rate of 96.9% due to the fact that it represents different characteristics of human gait. Finally, by fusing the three feature vectors, we found that the performance is increased to 97.4% because each set of features provides different characteristics of human gait, and together there is less negative conflict in terms of accuracy (see Figure 4.3). The last two results compare well with the state of the art.

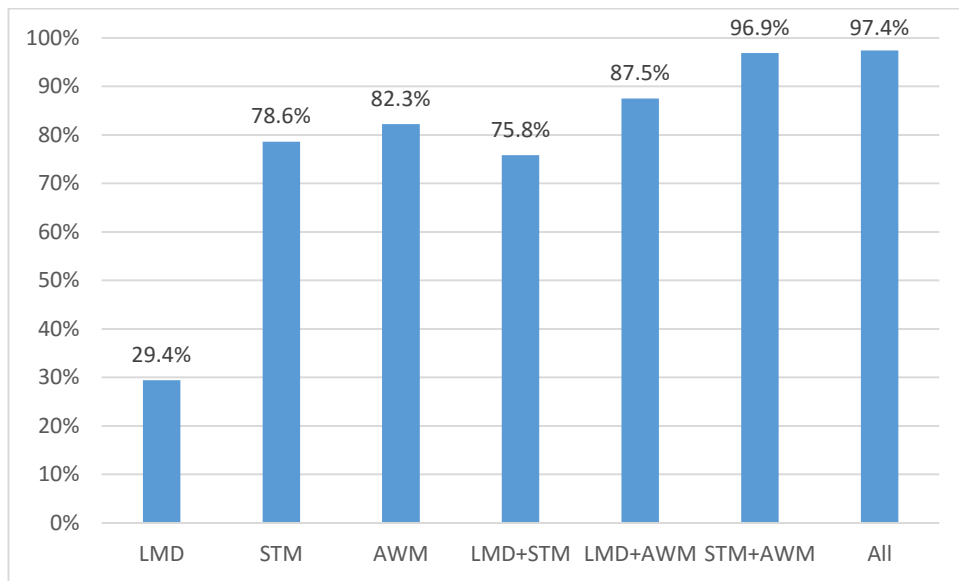


Figure 4-3: The Recognition Performance of NGR with LDA using each feature types LMD, STM and AWM and different combinations of the three feature sets. (The experiment is based on CASIA B gait database, using 124 subjects with 6 Nu, Gallery=4 Nu and Probe= 2 Nu for each subject).

In order to test whether the choice of LDA transformation is better option than that of using the PCA projection, we conducted experiments to test the performance of PCA-based transformation on the All-Features scheme and the dimensions are reduced into 14 dimensions, but we extended the experimental protocol to have different sample splitting between the gallery and the probe sets. In this experiment we also used two

different classifiers, the NN and LDC. Table 4.1 presents the recognition rate for all the combinations, these results show that LDA+NN does outperform the PCA based schemes. Moreover, the performance of this scheme does not deteriorate significantly when a smaller number of samples are placed in the gallery.

Table 4-1: Recognition Performance based on NGR using different number of samples in the gallery set. This experiment is based on CASIA B gait database, using 124 subjects and each subject has 6 Nu samples. (the reduced dimensions of the feature vector after applying PCA and LDA in this experiment are 14 and 6 respectively).

N. Gallery	N. Probe	NN+LDA %	LDC+LDA %	NN+PCA %	LDC+PCA %
4	2	97.4	97.3	81.4	96.0
3	3	96.6	95.9	78.5	95.3
2	4	94.4	94.4	74.9	93.2
1	5	75 based on (NN)			

Table 4.1 shows that the LDA features with NN and LDC outperform the two scheme that use PCA features with the NN classifier as well as the LDC classifier. Note that, in the last scenario, the gallery set has only one sample for each candidate, and hence the LDA cannot be defined in this case. Hence we only used the original features without any dimension reduction and classify these samples using the NN. This scheme achieves 75% recognition rate, comparing this result to that achieved based on all sets of feature in Figure 4-2, this result is less, due to having more sample in the previous experiment. We can also see that as the number of samples increases in the gallery set, the better the recognition results achieved by the system. Moreover the variance of the result in the first three scenarios is relatively small, when gallery = 2, 3, or 4 samples, and this demonstrates the level of performance of our system. This might be because of the more proper representation of the class's clusters when the number of samples is increased.

Reducing the dimensionality was done in order to achieve the ultimate dimension with the highest accuracy. To achieve this, we tried different LDA dimension reductions. Figure 4.4 shows part of our test which reflects the use of LDA from 10 to 1. This was applied to the All-feature scheme. The experiments show that the highest accuracy was achieved when the LDA set reduced the feature vector dimension to six.

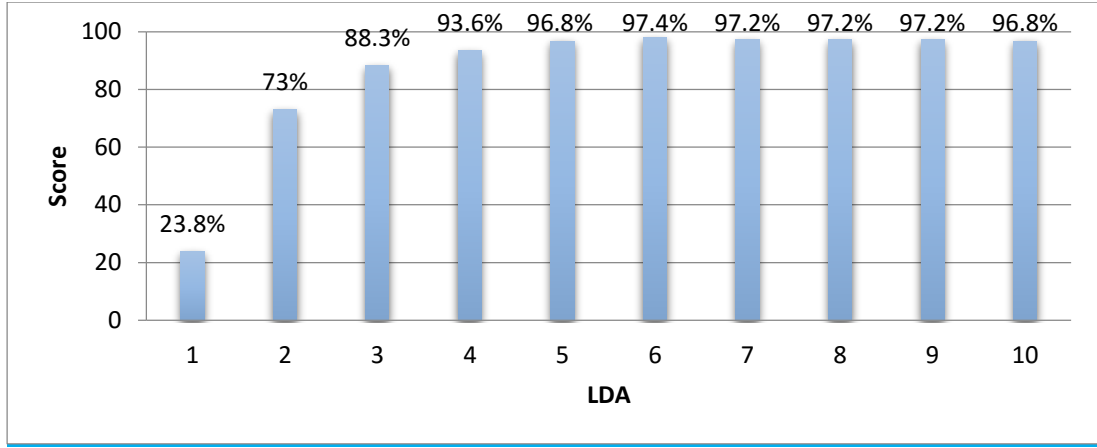


Figure 4-4: Recognition Performance of NGR based on fusing the three feature sets (LMD, STM and AWM) with different number of dimensions using LDA.

We close this section by comparing the performance of our method with that of the techniques proposed in (Martín-Félez, et al., 2011) and (Yu, et al., 2014) which used neutral gait recognition on the CASIA B dataset using leave-one-out cross validation. Nerveless (Martín-Félez, et al., 2011) and (Yu, et al., 2014) investigate other cross validation protocol, to compare our method with these two methods we implemented the proposed method based leave-one-out cross validation. The experimental results show that the proposed algorithm achieves 98.9% which is significantly better than 98.0 that provided by (Yu, et al., 2014) with $p=0.08$ (based on a chi square test) as shown in Table 4.2. The reasons that led to superior performance might be because of the proposed feature sets or due to the way of the pre-processing. Also applying wavelet transform followed by LDA as a dimension reduction technique makes our method computationally less expensive.

Table 4-2: Performance Comparison between proposed NGR method (using combination of the three feature sets LMD, STM and AWM) with other methods in the literature using leave- one- out cross validation, (based on CASIA B Gait database). RR is Recognition Rate.

Methods	N. subject	R.R.%
(Martín-Félez, et al., 2011)	124	97.0
(Yu, et al., 2014)	100	98.0
Our Method	124	98.9

4.2 The Reliability of NGR decisions

The reliability of decision made by any system is usually very important in the process of identification. This is to do with the fact that Identification/classification decisions are often based on distances/similarity values for the probe sample against all the different subjects (classes). Once a representative of a class is identified as satisfying the matching criteria no attention is paid to other gallery templates that may be of different classes and achieve similar scores or slightly different score. Such a scenario, would raise serious issues on reliability of the decision. Many identification researches report high identification accuracies without presenting any measure of reliability of the individual decisions. Therefore the space and the distribution of these scores, when a probe sample is tested against all gallery templates, has an influence on reliability of the decision. Moreover, there Nearest Neighbour classifier always returns the class of the nearest neighbour even if the probe sample is not enrolled in the biometric database. Here we shall introduce and investigate a simple reliability measure which assumes that all probes are enrolled (see Figure 4.5). Different measures can be defined to cover other cases when dealing with persons not enrolled, but this would require an in-depth analysis that is beyond the remit of this thesis.

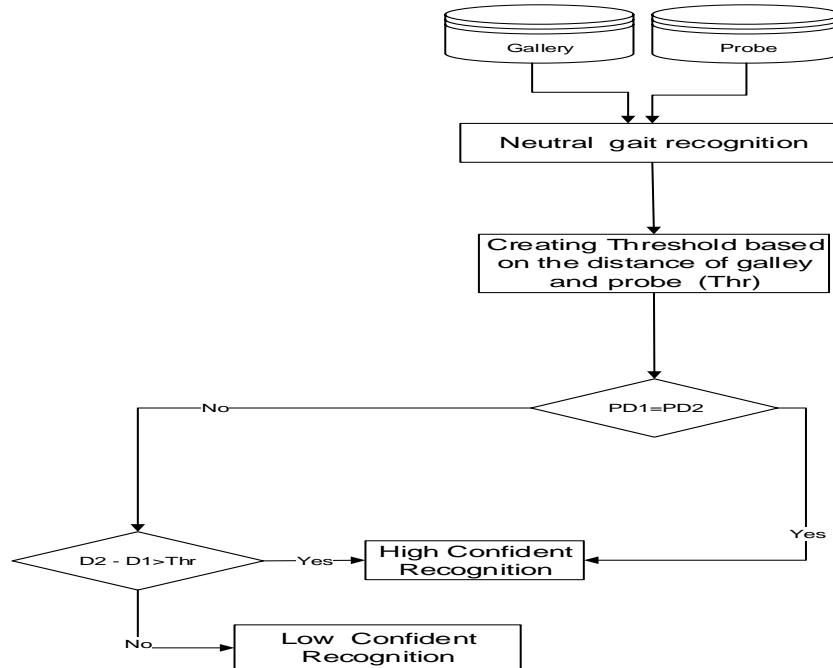


Figure 4-5: Overview of the reliability in the proposed NGR method. Where D1 is the first nearest distance; D2 is the second nearest distance, Thr. is the created threshold and PD1 and PD2 are the first and second nearest person from the probe sample respectively.

The Simple Reliability measure

First assume that a probe p was recognised as belonging to a class c because its nearest template sample x_1 is from class c . We also assume that the system has K different classes.

1. If the next nearest template x_2 also belongs to c , then we return a highly reliable decision,
2. Else: calculate the differences between the distances of the probe to the two nearest templates, i.e. $d = D(p, x_2) - D(p, x_1)$. The reliability of the decision is now dependent on how large is d relative to the difference $d_{\max} = D(p, x_{\max}) - D(p, x_1)$, where x_{\max} is the furthest away template from the probe.

If $d > (d_{\max} / K)$ then return a Highly Reliable (HR) decision. Otherwise we say that the decision is of low reliability.

The HR measure is defined as:

$$HR = 100 \times \frac{\text{No. of highly reliable decisions}}{\text{Total number of tests}}. \quad (4-1)$$

To implement this method we used CASIA B gait database that contains 124 subjects and each subject has six samples. Gallery and probe set will be created based on different scenarios in relation to the number of samples in the gallery and the number of probe samples per subject.

Table 4.3, present the outcome of the previous experiment on gait recognition but with the HR values for each of the experimental scenarios. These results show that we can be highly confident that the decisions made by our NGR gait scheme are highly reliable in more than 90% of the times.

Table 4-3: Recognition Performance of NGR and the reliability of its decisions, and presenting the result using different scenarios of different number of samples used in the gallery and probe set for each candidate (this experiment is based on CASIA B Gait database 124 subjects with 6 Nu gait sequences each, using).

No. of Gallery	No of Probes	Recognition Rate %	High reliability %	Low reliability %
2	4	94.4	91.8	2.6
3	3	96.6	94.1	2.5
4	2	97.4	95.0	2.4

4.3 Gender classification from neutral gait sequences

After analysing the experimental results, we conclude that when fewer samples are included in the gallery, more female candidates were misrecognised and confused with male candidates in the system, while this is not the case for male candidates (see Table 4.4).

Table 4-4: The percentage of confusing female with male and also male with female using Neutral Gait Recognition, (CASIA B gait database).

G	P	Female/Male (%)	Male/Female (%)
1	5	20.43	2.02
2	4	2.15	0.34
3	3	1.34	0.01
4	2	0.27	0

These observations motivates the development of a gender classification based on neutral gait sequences the developed scheme, that would be presented here, will exploit each of the three different types of features (STM, LMD and AWM) singularly as well as combined. For testing the performance of our method we used CASIA B gait database, and we randomly select equal subsets of males and females to run the experiment repeatedly 30 times to cover the entire subjects in the database. Note that there is an imbalance gender representation in the database. For classification we use the k-Nearest Neighbours (k-NN). Moreover due to having two classes in this method we also used Support Vector Machine (SVM), originally designed for binary classification, and commonly employed for gender classification.

4.3.1 Gender Classification Scheme (GCS)

After extracting the silhouettes, three different feature vectors, described in chapter 3, are extracted for use in this method: STM feature vector deal with Spatio-Temporal distance; LMD feature representing the variation of energy in legs between wavelet subband LL1 of consecutive frames during one gait cycle measured by Hamming distance; and the AWM feature vector representing the distribution of the LL3 wavelet subband. All three feature vectors are extracted separately, and then fused using three types of fusion: feature fusion, weighted score fusion, and decision voting fusion. In this system we used two different classification methods: k-Nearest Neighbour (k-NN)

and Support Vector Machines (SVM) with different kinds of fusion (see block diagram in Figure 4.6).

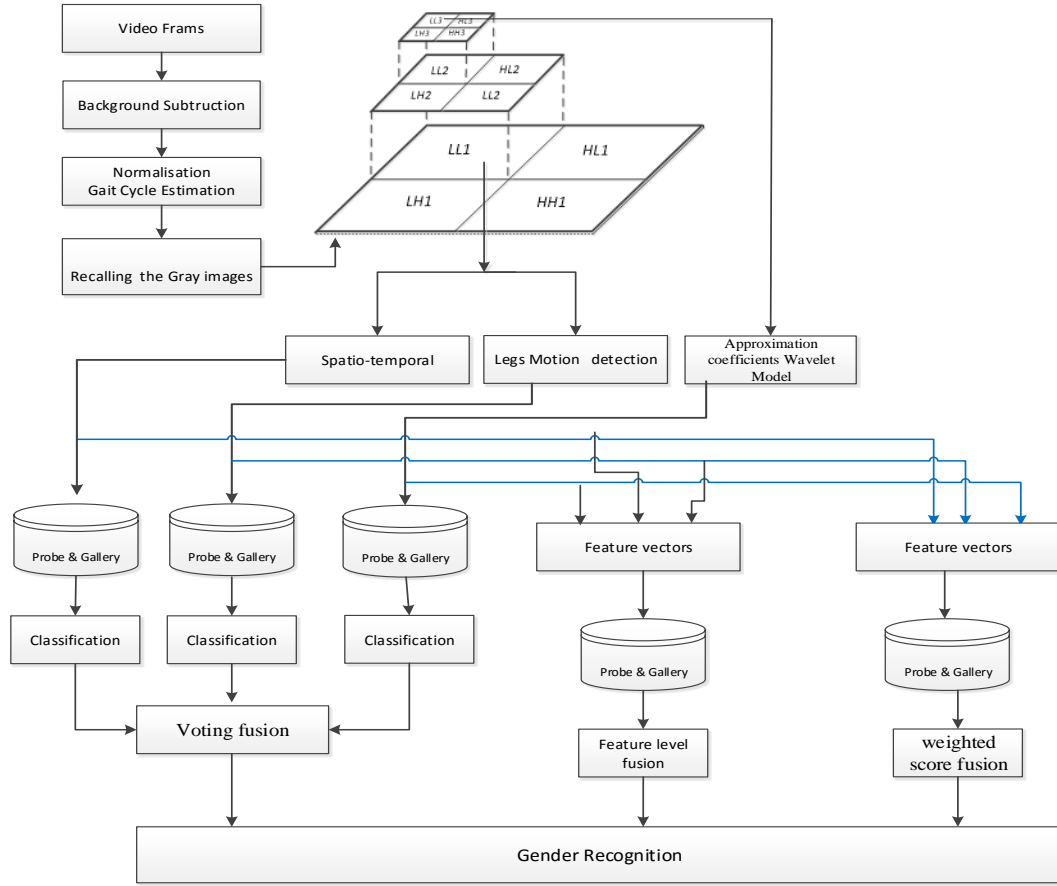


Figure 4-6: Overview of Gender Classification Scheme based on neutral gait sequences (In the weighted score fusion the individual weights were based on the ratio of its accuracy rate to the sum of accuracy rates for the 3 schemes).

4.3.2 Experiment Setup

To test the performance of the above GCS method we use the CASIA B gait database. Unfortunately, this database is imbalanced in terms of number of male and female participants (93 males and 31 females). To avoid possible bias in the accuracy rates, we randomly select subsets of 25 males and 25 females from the main sets (each subject from male and female has 6 records) and repeat the experiment 30 times, and calculate the average recognition rate. In this experiment, 10-fold Cross Validation (CV) is used in conjunction with the two classification method (SVM and k-NN). The kernel function used for the SVM is Linear with SMO method for optimization. In the following subsection we take $k=1$, i.e. we use the NN classifier, but we shall also present and compare accuracy rates for $k=1, 3$, and 5 .

4.3.3 Experimental Results

The results of the first experiment where k-NN, $k=1$, was used is presented in figure 4.7, where the average accuracy rate and standard deviation are calculated over the 30 repetitions. The STM-only based scheme outperform the other single-feature schemes achieving 94.2%. This was because this set of features deals with different parts of the human body such as feet, hands, shoulders and the height of the candidate, and each one of these features are affected by the differences in the whole body structure of male and female. This also explains why the AWM is the second best performing by a large margin when compared to the LMD scheme which is related to legs only. Fusing AWM and STM at the feature level has achieved 96% accuracy outperforming the fusion of other pairs of features. The fusion of the three features (STM, AWM and LMD) at the feature level has not exceeded the 96% accuracy, but a weighted score fusion of the 3 single feature schemes has improved the accuracy and provide 96.5%. The individual weights were based on the ratio of its accuracy rate to the sum of accuracy rates for the 3 schemes. On the other hand majority voting for the decision-based fusion reduced the accuracy to well below the STM scheme.

From these experiments we can see that the standard deviation of accuracy achieved by the individual feature vector cases (AWM, LMD or STM) are relatively small (< 3) over the 30 experiments indicating that the average accuracy rates are indicative of the actual rate for these schemes. In the cases where we fuse all the 3 types of features, with or without weighting, the situation are even better as a result of std being < 1 .

In table 4.5, we present the accuracy rates for the above single and fused schemes when use kNN classifier with $k=1, 3$ and 5 . These result shot that except for the LMD feature the result for $k=1$ outperform the schemes that uses $k=3$ which in turn outperform the schemes when $k=5$, i.e. the performance of these schemes deteriorate as k increases.

The results shown in Figure 4.8 relate to the second part of the experiment when we use the of SVM classifier. Based on this experiment, we conclude that fusing STM and AWM achieves the best rate of 95% as compared to other combinations. However, except for the case of using LMD, the results show that the SVM was outperformed by the k-NN, with $k=1$. This last conclusion is illustrated better in Figure 4.9.

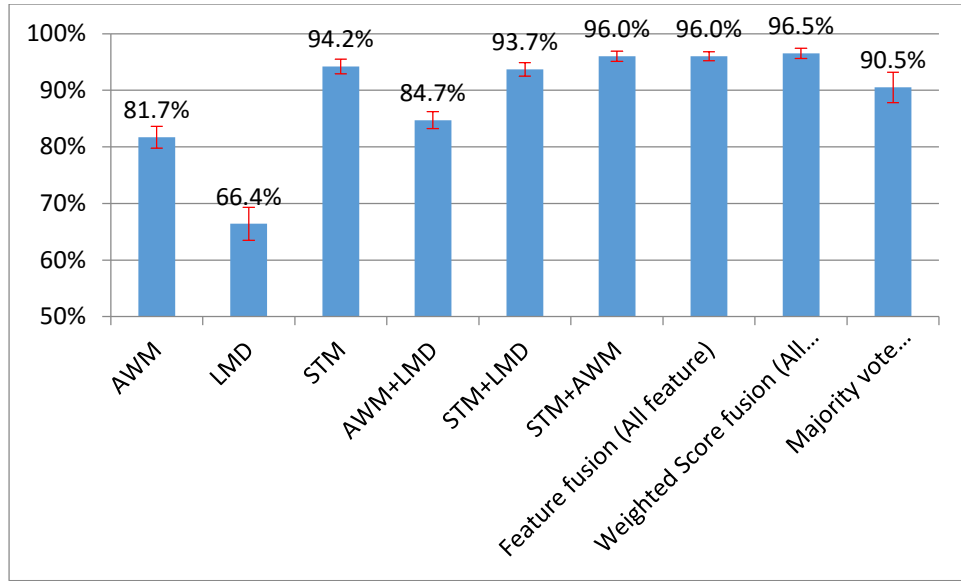


Figure 4-7: The Average and standard deviation of the GCS. using the three different features (AWM, LMD and STM) separately and in combination using k-NN (k=1).

Table 4-5: The performance of Gender Classification Scheme using kNN with k=1, 3, and 5

	K=1(%)	K=3(%)	k=5(%)
AWM	81.7	80.9	80.6
LMD	66.4	67.4	68.7
STM	94.2	93.2	92.0
AWM+LMD	84.7	81.9	81.1
STM+LMD	93.7	93.1	92.1
STM+AWM	96.0	95.3	94.0
Feature fusion (All feature)	96.0	95.6	94.5
Weighted Score fusion (All feature)	96.5	95.2	94.4
Majority vote fusion (All feature)	90.5	89.6	89.0

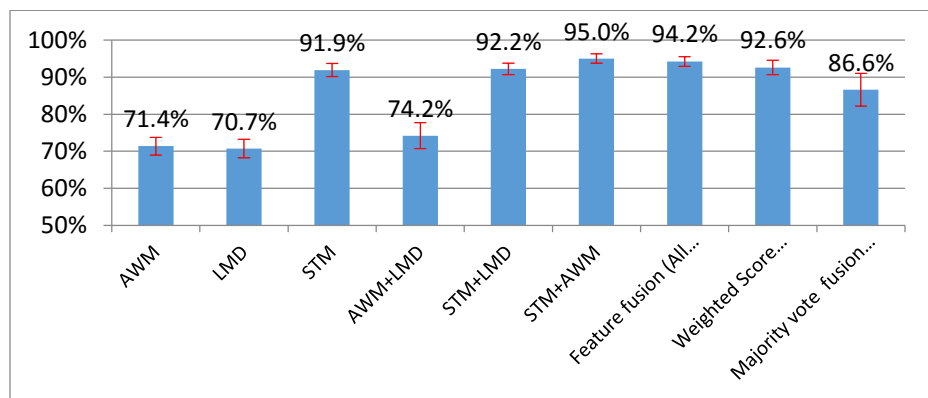


Figure 4-8: The Average and standard deviation of the Gender Classification Scheme (GCS). Using three different features separately and in combination using SVM.

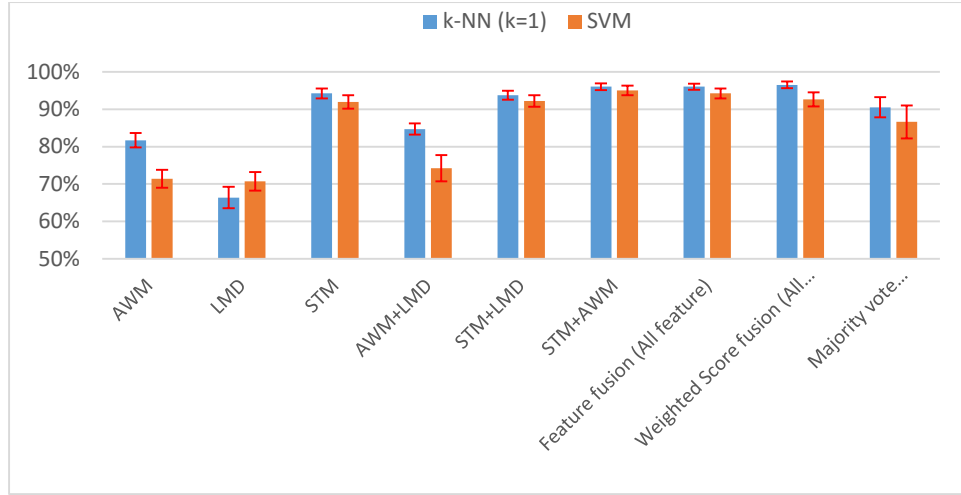


Figure 4-9: Performance Comparison result of using k-NN and SVM with Gender Classification Scheme (based on CASIA B Gait database) .

In order to determine the effect of varying the number of folds in the cross validation, we repeated the best performing scheme (the weighted score fusion of all the 3 features) but with different cross validation protocol using k-NN with $k=1$. The results of these experiments, shown in figure 4.10, demonstrate that the best recognition is achieved with the 10-fold cross validation and the performance decline as the number of folds decreases.

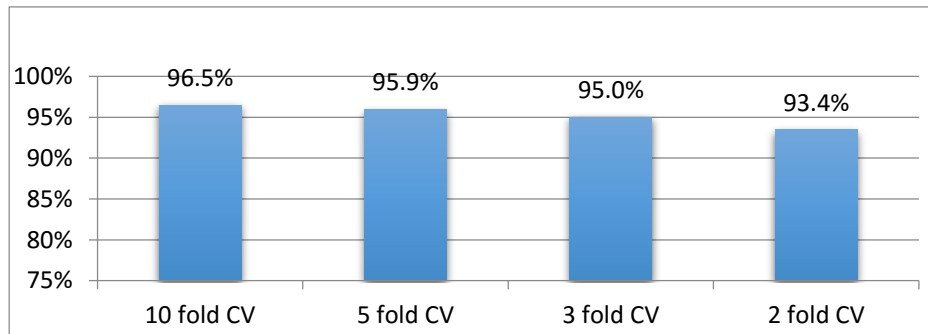


Figure 4-10: The effect of varying the number of folds in the cross validation on the Recognition Performance, based on k-NN ($k=1$) using Weighted Score fusion of all feature sets.

Finally, we compare the performance of our best scheme (i.e. weighted score fusion of all features), using the k-NN, ($k=1$) classifier with related existing schemes published in the literature, some of which may use different classifiers. Table 4.6 illustrates our results together with the results achieved by these methods. For each method the table show the number of male and females used, the cross validation

protocol adopted, and the number of repeated experiments. The results demonstrate that our scheme outperform all the other methods, achieving 96.5% on average.

Table 4-6: Performance Comparison of GCS with other methods using CASIA B Gait database, CR is Classification Rate.

Method	Dataset (Male/Female)	Cross Validation	CR %	N. of sets
(Arai & Asmara, 2011)	31/ 31	10	92.9	1
(Chang & Wu, 2010)	25 /25	5	92.3	1
(Martin-Félez, et al., 2010 'A')	93 / 31	10	91.8	10
(Martin-Felez, et al., 2010 'B')	93 / 31	10	94.7	10
(Yu, et al., 2009)	31 / 31	31	95.9	1
(T. Anitha, February 2013)	31 / 31	31	93.3	1
(Arai & Asmara, 2012)	31 / 31	10	97.6	1
GCS	93/ 31	10	96.5	30

4.3.4 Reliability of GCS Decisions

In order to measure the reliability of decisions made by GCS scheme we follow the same simple procedure described in section 4.2 including the same threshold that was worked out as the ration of the max-min distances from the probe by the number of subjects. Recall the threshold is only applicable if the two nearest neighbours to a probe are of different gender. To choose the threshold for the GCS scheme, which only have 2 classes (Male or Female), by taking half of the (max-min) distances is not realistic because it would assume that the persons within each of these classes are very well separated. This is based on the fact that the experiments of the NGR schemes revealed that the spatial locations of the two genders are overlapping. The experimental result show that we can be highly confident that the decisions made by our GCS gait scheme are highly reliable in 95.5% of the times.

Gender classification can be applied in various applications such as smart surveillance systems and demographic studies systems. In addition, gender can be used in gait recognition systems in order to reduce the time needed for searching a gait database based on one gender only, and improving the accuracy. In the next section, in order to consider the final advantage of gait-based gender classification, we combine gait recognition with gender classification in a neutral gait sequence, aiming to enhance the gait recognition performance.

4.4 The 2-level Neutral Gait Recognition (NGR2)

In this method we try to benefit from gender classification, presented above, to improve the performance of NGR scheme in terms of gait recognition. Following on from Table 4.7, regarding the confusion of females and males, we collected similar data for two of the high performing gait recognition schemes: STM +AWM (96.9), and STM+LMD+AWM (97.4%). Table 4.7 present the results of confusing female with male in NGR using FV1 and FV2 separately, where FV1 refers to (STM+LMD+AWM) and FV2 refers to STM+AWM. The table show that FV2 is better gait recognition performance for female candidates, compared to FV1.

Table 4-7: The percentage of confusing female with male in Nuetral Gait Recognition (NGR) based on FV1 and FV2.

G	P	Female/Male using FV1 (%)	Female/Male using FV2 (%)
1	5	20.43	0.83
2	4	2.15	0.19
3	3	1.34	0.075
4	2	0.27	0

4.4.1 The NGR2 Algorithm

Using the above observations, one can develop 2-level neutral gate recognition. One may start at the first level to determine the gender of the newly observed person, and then run the NGR scheme with FV1 (or FV2) when male (or a female) was declared at the first level. However, in this way level 1 decision may prejudice the decision of the second level. Hence, we opt for running both NGR and GCS in parallel at the first level, and at the second level we first check if the gender labels from the procedure are identical or not, (see Figure 4.11). Let G1 be the gender label of the person identified using the NGR with FV1, and G2 be the output label from the GCS FV2. If $G1=G2$ then compare the distances between the probe and the two identified persons. Note that, in this case the two schemes may identify different persons, of the same gender. If G1-person is the nearest then return G1-person, otherwise run NGR but using FV2. On the other hand if $G1 \neq G2$, and G1 is male then run NGR using FV2 otherwise return the G1-person as identified in Step1.

We shall next test the performance of NGR2 and compare with that of NGR, for different protocols using CASIA B database.

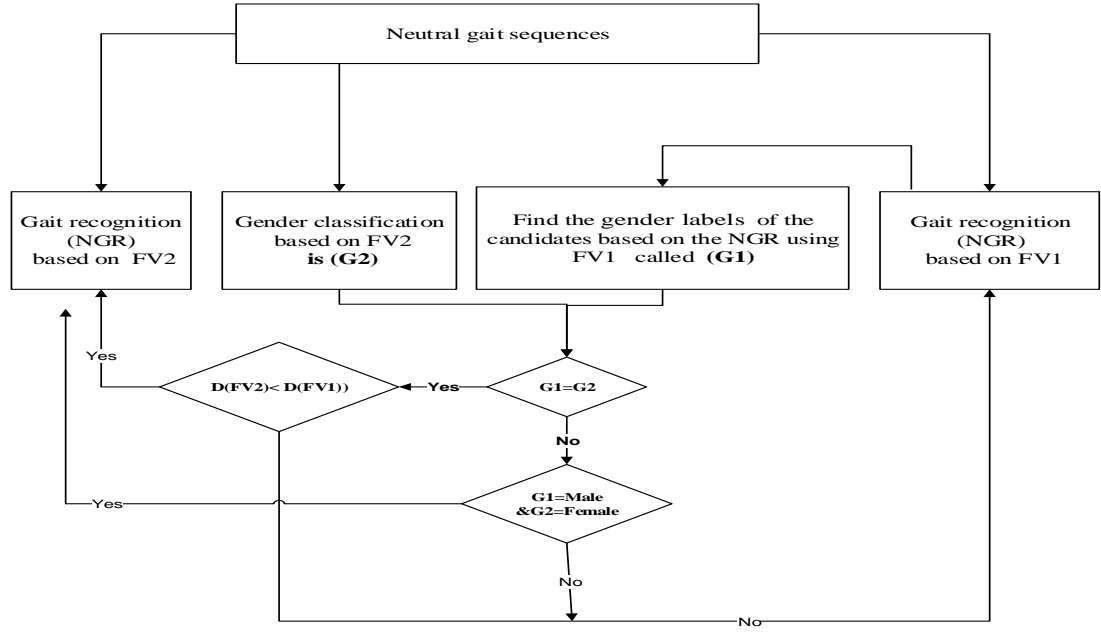


Figure 4-11: Overview model for combining gender classification with gait recognition. $D(FV1)$ and $D(FV2)$ are the distance of recognized person from NGR using FV1 and FV2 respectively.

4.4.2 Experimental Results

Table 4-8 displays the result of the performance testing experiments of the NGR2 scheme in comparisons to that of NGR using various samples selection scenarios. In this method, the gender is trained based on 20 female and 20 male samples that are randomly selected from the gallery of the NGR scheme, and then k-NN ($k=3$) is used to classify the gender. The results demonstrate clearly that NGR2 outperforms the NGR in most of the scenarios, albeit marginally, except in the case when the gallery contains one video per person. This validates our unstated hypothesis that there is a reasonable chance of enhancing gait recognition performance when the gender of the probe is known in advance. Moreover, we note that interleaving the use of two different gait-related schemes that are known for their high performances can yield improved results,

Table 4-8: Recognition Performance of Neutral Gait Recognition with and without gender classification, (NGR is gait recognition without gender classification and NGR2 is gait recognition with gender classification).

N. gallery	N. Probe	NGR (%)	NGR2 (%)
1	5	75	78.3
2	4	94.4	95.2
3	3	96.6	96.9
4	2	97.4	97.4

4.5 Conclusions

This chapter was devoted to investigate neutral gait sequences for use in gait recognition. Three different sets of wavelet based features were extracted from different subband levels and used as gait signatures: the STM (from LL1), the LMD (from LL1), and the AWM (from LL3). Although these different sets are not of very high-dimension, we applied the known dimension reduction schemes of LDA and PCA in order to extract lower dimensional feature vectors that involve less redundant attributes.

Different recognition schemes were designed using each feature sets separately as well as all combinations of feature level fusion where these signatures are concatenated. The performances of these schemes were tested on the CASIA B gait database using two different classifiers: the LDC and the NN. The experimental results demonstrated that, when no dimension reduction was applied, the single schemes of STM and AWM and their fusion achieved relatively good accuracy rates, while the performance of the LMD scheme was somewhat disappointing and its fusion with STM or with AWM led to some modest improvement but when fused with the STM+AWM the performance deteriorated. The effect of using the LDA dimension reduction resulted in much improved accuracy with best accuracy rate of 97.4 achieved with the fusion of the 3 feature sets. The performance of the LMD scheme did improve with dimension reduction. To show the efficiency of the proposed method we investigated a simple reliability method to be used for NGR decisions. After analysing the experimental results of NGR, we notes that misrecognized female candidates are mostly confused with male candidates in the gallery set, to overcome this problem we investigated gender classification based on gait signature. In the gender classification method we used the same three feature vector (STM, LMD and AWM) were used for gait recognition based on the same database (CASIA B). In this approaches we used different types of fusion; feature level, decision level, and weighted score fusion. SVM and kNN are used as a classification method. The experimental results show that the proposed method outperforms existing methods. By integrating gender classification in gait recognition in the 2-level Neutral Gait Recognition (NGR2), we established that knowledge of gender provides better performance as compared to NGR.

There are many factors that influence human gait recognition. Some of these are internal factors that affect the natural gait such as pregnancy and aging, while the other

factors - the so-called external factors - are mostly aspects that impose a challenge when it comes to recognition such as clothes and carrying conditions. To address some of these challenges we will address the external factors by focusing on clothing and object carrying in the next chapter.

Chapter Five

Unrestricted Gait Recognition

In chapter 4 we focused on gait recognition/identification using neutral gait sequences. In this chapter we will expand the work to investigate the problem of identifying individuals based on unrestricted gait sequences, i.e. without restriction and regardless whether the person is in neutral state or carrying bag and/or wearing coat. The last case is referred to as gait in non-neutral circumstances. We first extend the use of the NGR scheme for the unrestricted scenario to initiate our investigation into UGR by testing the possibility of using the same, or a modified version, of the 3 sets of features (STM, LMD and AWM) for gait recognition. We then investigate the use of 3 other types of features for general gait recognition. These sets of features are the Gait Energy, the Gait Entropy and their fusion. Although each of these different new feature vectors improves gait recognition by considerable amounts, further investigations will be conducted to improve these results. We shall follow the approach adopted in chapter 4, and attempt to develop a non-neutral gait sequence case detection (GSCD) algorithm and gender classification to be incorporated into an advanced gait recognition scheme for the unrestricted scenario. We shall demonstrate that this approach yield improved performance.

5.1 Extending NGR to Unrestricted Gait Recognition.

The unrestricted gait sequence scenario is concerned with recognising a person from his/her gait sequence in 3 different cases: Neutral gait sequences (Nu), Carrying Bag sequences (CB) and Coat Wearing sequences (CW). Person recognition in the unrestricted scenario is very challenging due to the fact in the non-neutral cases of carrying bags and coat wearing, there is no standard that can be applied to type of bags carried or coats worn. Having found, in chapter 4, that the fused 3 wavelet based features (STM+LMD+AWM) achieved excellent for the NGR (>97% accuracy) it is natural to test the performance of that scheme for the non-neutral gait sequences. Due to the inability to standardise bag sizes and coats, we conducted an gait recognition experiment whereby the gallery consists the first 4 neutral gait feature vector samples

per person the probe set consisting of the last 2 samples per person for each of the conditions (i.e. 2 neutral samples, 2 CB samples and 2 CW samples).

Figure 5.1 show that the 3-component fused gait feature vectors that in chapter 4 was shown to achieve good result for NGR, does not have acceptable discriminating power when used for recognising gaits under the CB and CW conditions. Since the gallery in these experiments we made up entirely of neutral templates, the accuracy remained high for the neutral test samples, but for the non-neutral cases the results deteriorated significantly with only 25% for the CW case with an overall accuracy of only 64.5%. This experiment, shows the need to modify these features in some way or consider other new features that can be extracted from the gait sequences for improved accuracy of UGR. In the rest of this section we shall investigate the first option.

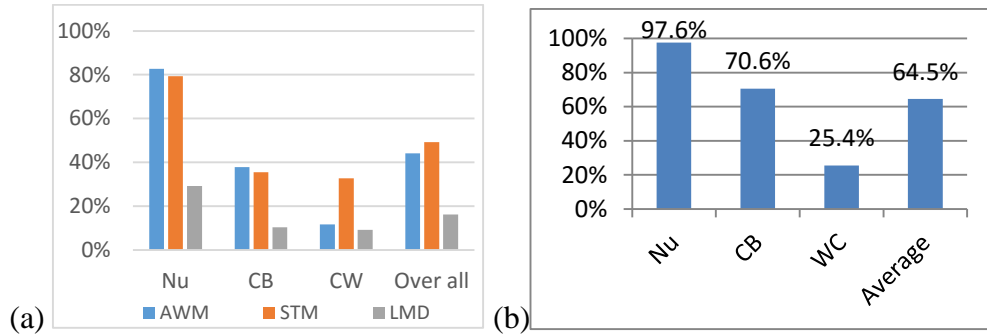


Figure 5-1: Recognition Performance of using NGR feature vector for UGR (Figure 5-1 (a) is the result of using each feature set separately (AWM, STM, and LMD) for each of Nu, CB, and CW. Figure 5-1 (b) is the result of the combination of (AWM, STM, and LMD) for each of Nu, CB, and CW.

The three features are extracted from the low frequency wavelet subbands at level 1 and level 3 (i.e. LL1 for STM and LMD and LL3 for AWM). The fact that wavelet transforms are multi-resolution analysis tools that decompose images into multiple ranges of low and high frequency waveforms at different scales raises and the high frequency subbands contain information about significant image features in an efficient manner, raises the possibility of extracting similar features from the non-LL subbands. The LL1-subband is a good approximation of the spatial domain of a wavelet decomposed image, while the non-LL subbands at level 1 represent the lost details from the original image as a result of the approximation wavelet procedure. The STM features are related to distances between different parts of the body and LMD models differences in areas of the leg during movement. So it is natural to extract

those two sets of features from LL1, because the non-LL subbands provide little or no chance of determining the necessary distances areas across frames of gait sequences. Moreover, the results in figure 5.1(b) and those in chapter 4 illustrate that modifying LMD can only yield marginal effect, and among the other two the overall performance of the AWM seems to provide a more promising choice for improvement. The AWM features represent statistical parameters of the distribution of LL3 coefficient, and these distributions surly depend on the variation in type of bags carried, the way the bags are carried and the coats worn. It is a well-known fact that the for any image the distribution of the coefficients in any non-LL wavelet subband is always a Laplacian distribution with 0 mean (also known as generalised Normal distributions) regardless of the image or the wavelet filter (Al-Jawad, 2009). Therefore it is more sensible to consider extracting the statistical parameters, used in AWM, but in the non-LL subbands at all levels up to and including 3. Hence we define the Detail coefficients Wavelet Model (DWM) to the feature vector consisting of 2 statistical parameters obtained from the sequence the non-LL subbands at level 1, 2, and 3. In each non-LL subband these two parameters are the *mean* and *std* of the sequence of the standard deviation of the given subband coefficients obtained from the sequence of frames in the gait cycle. The precise definition of DWM is as follows:

Let $\{F_1, F_2, \dots, F_k\}$ be the sequence of frames in a gait cycle, and $\{WF_1, WF_2, \dots, WF_k\}$ be their wavelet decomposition to level 3. For each wavelet frame WF_i , arrange the non-LL subbands in $S_{i,1}$, the order:

$$\{S_{i,1}, S_{i,2}, \dots, S_{i,9}\} = \{LH3i, HL3i, HH3i, LH2i, HL2i, HH2i, LH1i, HL1i, HH1i\}$$

For each i, j , let $std_{i,j}$ = be the standard deviation of the coefficients in $S_{i,j}$, and let

$$\mu_j = \text{mean} \{std_{1,j}, std_{2,j}, \dots, std_{9,j}\}, \text{ and } \sigma_j = \text{standard deviation} \{std_{1,j}, std_{2,j}, \dots, std_{9,j}\}$$

For the given cycle, DWM^p feature for the gait cycle in the upper or lower part of the body is defined as:

$$DWM^p = \{\mu_1, \sigma_1; \mu_2, \sigma_2; \dots, \mu_k, \sigma_k\}.$$

And the whole body feature is defined as:

$$DWM = \{DWM^L, DWM^U\}.$$

5.1.1 The modified Unrestricted Gait Recognition (UGR) algorithm.

In this method we fuse the set of wavelet feature vectors {STM, LMD, AWM, and DWM) to represent a gait features. These features are extracted from the pre-processed sequence of frames in the gait cycles. Before testing the performance of this scheme, we recall that the purpose of introducing these new sets of features was to improve the recognition rate in the non-neutral case. Since the lower part of the body remains more visible/measurable variant in the walking style of the person, the question arises as to whether we need to extract these features from the whole body or from lower part only. Moreover, upper body region provides information about body shape and stance and also represents more constant part of human body, while the lower body part offers the dynamic feature and the whole body represents both static and dynamic information. To answer this question we decided to test the performance of the scheme by extracting the features once from the whole body and once from the lower part. The lower part of the body is determined by the golden ratio proportion (Herman, 2007) .

5.1.2 Experimental Results

To test the performance of the proposed method we use the CASIA B gait database; this database includes 124 subjects. For each subject there are ten walking sequences consisting of six Nu gait sequences (where the subject does not carry a bag or wear a coat), two CB gait sequences and two CW sequences. As before, the selected viewing direction in our proposed method is 90 degrees (i.e., side view). In what follows S1 is the fused feature vector (STM+LMD+AWM+DWM) extracted from the whole body, while S2 is based on the same fused features but the last two components are extracted from lower body part. We also apply the LDA method as a dimension reduction method to reduce the feature vector dimensions for both S1 and S2, the Nearest Neighbour (NN) is used as the classifier. The experimental results, shown in Figure 5.2 demonstrates that with both sets (S1 and (S2) of features the performance of the UGR is significantly better than when we extended the NGR to non-neutral cases. However, the accuracy of S1-based UGR on the CW case is still disappointingly low, while the accuracy of the S2-based UGR on the CW improves remarkably to 81%.

We explain this by the fact that wearing a coat results in covering nearly $\frac{2}{3}$ of the body and thereby reducing variation in the S1 features extracted from the covered part, during the gait cycle. Moreover, for the S2-based UGR the AWM and DWM features are extracted from the lower body part which is mostly not covered by the coat. However, the S1-based UGR outperforms the S2-based UGR for the Nu and the CB cases.

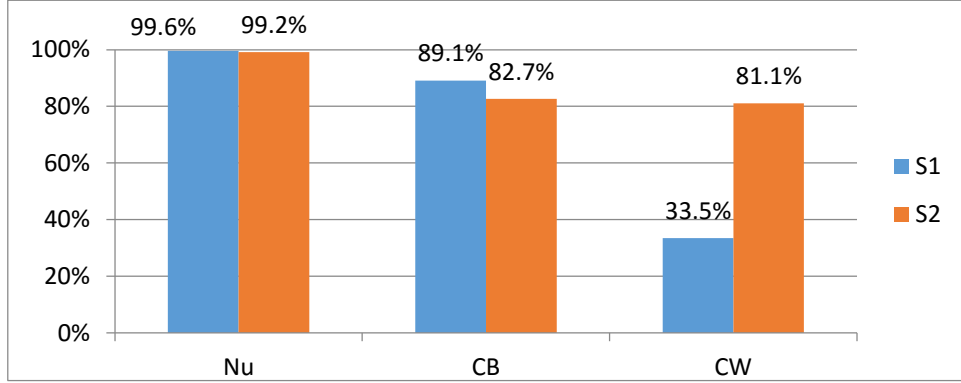


Figure 5-2: Recognition Performance of the Unrestricted Gait Recognition for S1 and S2 feature vectors separately. S1 is (STM+LMD+AWM+DWM), while S2 is {STM+LMD+ lower body part (AWM+DWM)}.

Although carrying bags does have an effect on variation in whole body movement but not to the point where variation cannot be detected as in the case of CW. By comparing CB with CW we conclude that the effect of the bag on the human body is less severe on the persons walking style than the effect of wearing a coat, knowing that the gallery contains only a neutral gait sequence. But this may not be the case if the carried back is bulky and/or heavy, which is not represented in the CASIA B database.

In the above experiments, we have the gallery consisting of only 4 Nu samples, and one can ask what happen if we consider other configurations. We conducted some experiments to test this kind of arrangements. Table 5.1, shows the accuracy rates of the S2-based UGR for a limited number of combinations of gallery to probe ratios. The first 3 columns are the result in Figure 5.2. Although using other combinations whereby gallery could contain a mix of sample, may be useful, in this case the fact that there are more samples of Nu that could overwhelm the experiment. There we will not do that here.

Table 5-1: Recognition Performance of using UGR for a limited number of combinations of gallery to probe ratios using S2 (STM+LMD+ lower body part (AWM+DWM)) feature vector.

P-G	Nu-Nu	CB-Nu	CW-Nu	CW-CW	CB-CB	CW-CB	CB-CW	Ave
Gallery(G)	4(Nu)	4(Nu)	4(Nu)	1(CW)	1(CB)	2(CB)	2(CW)	
Probe(P)	2(NU)	2(CB)	2(CW)	1(CW)	1(CB)	2(CW)	2(CB)	
Accuracy	99.2%	82.7%	81.1%	62.1%	54.8%	68.7%	65.7%	73.5%

The experimental result that present in Table 5.1 shows that for the non-neutral cases the accuracy rates goes down significantly when CW and CB were used in the gallery as opposed to their performance when Nu sample were used for the gallery. Recall that in Chapter 4, knowledge of the walking person gender improved the NGR performance. Here we investigate, in the next section, the problem of detecting the case of the walking person, i.e. whether the walking person is carrying a bag, wearing a coat, or neither. This could help achieving improved results for the non-neutral case than those presented in Figure 5.2, by selecting the best features among the S1 and S2.

5.2 Gait Sequence Case Detection (GSCD)

Detecting weather a person walking into the camera view is wearing a coat, carrying a bag, or neither is another pattern recognition task which requires a knowledge on the effect of these extra item on the person's gait. We know that the non-neutral cases have an impact on the movement of the shoulder and of the pair of hands. Hence, the most relevant features must be extracted from those two parts of the body. Hence we use two sets of features. The first set is constructed from the distance between hands and shoulders during one gait cycle which are usually extracted as a part of STM. Since shoulders and hands are in the upper part of the body, then the second suitable feature set is the AWM extracted from upper body part only. The two sets of feature were fused using feature fusion method to be used as one feature vector.

The GSCD algorithm is therefore concatenate the STM and AWM features extracted from the upper part of the body. Again we test the performance of the GSCD on the CASIA B gait database. The adopted experimental protocol, uses 20% of the entire dataset as the gallery with balanced choices of 2 samples per subject in each of the 3 cases: Nu, CB and CW. The remaining 80% of the data form the probe set to be used for testing. We repeated the experiments 5 times by random selection of samples and the average of the results were presented as a performance of the method. Figure 5.3

shows the results of using the LDC and the k-NN ($k=7$) classifiers separately. It demonstrates that the LDC yields significantly better detection of the non-neutral cases than the kNN. But the kNN is slightly better in detecting the neutral case. The extremely high success detection of the CW case by both classifiers is very important for incorporating GSCD into UGR since the CW is the case where UGR schemes are not performing as well as in the other cases.

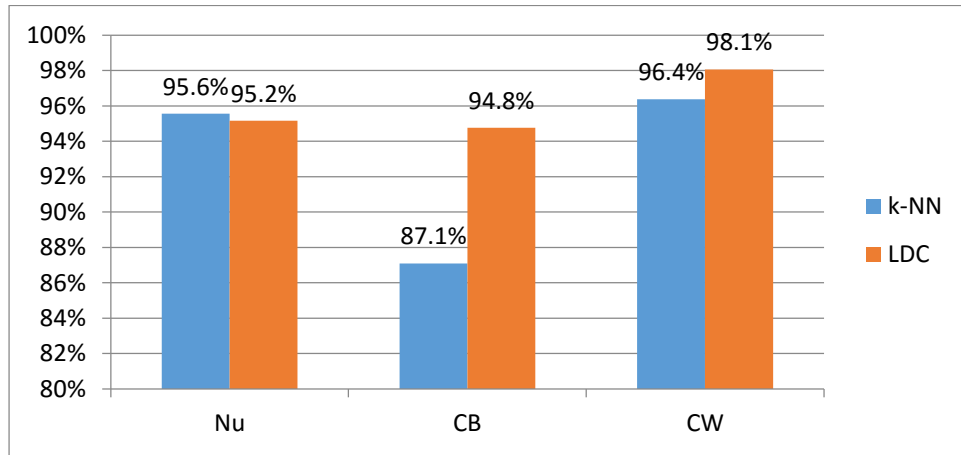


Figure 5-3: Recognition Performance of Gait Sequence Case Detection using k-NN and LDC based on CASIA B gait database.

5.3 The GSCD-UGR scheme

Here we modify the UGR scheme by incorporating the GSCD algorithm in a very simple manner, using the results from Figure 5.2 which shows that the S2 feature vector yield extremely significant accuracy when used to recognize persons wearing coats. Hence, this algorithm first uses GSCD to determine the case and uses S2-based UGR scheme if CW was detected, otherwise uses the S1-based UGR scheme. Figure 5.4 displays the performance of this modified version of UGR when test on the whole of the experimental CASIA B gait database. Note that the detection accuracy of the GSCD is based on using only 20% of CASIA B. The experimental results on the whole database confirm the best results obtained in Figure 5.2 for each case.

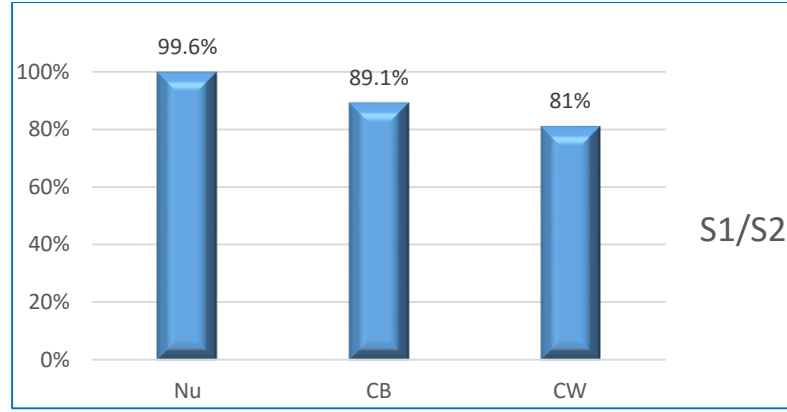


Figure 5-4: Recognition Performance of incorporating Unrestricted Gait Recognition UGR with Gait Sequence Case Detection (GSCD).

The above results show the importance of know the case in advance. However, misdetection in the proposed GSCD method does not mean that there is no chance for recognising the person. For example if the CW gait sequence was detected as a Nu case then the person may still be correctly recognised due to the fact that UGR recognise the person with an accuracy of 99.6%. This may be also influenced by the fact that the gallery was represented by 4 Nu samples which is twice as many as the number of tested samples in each of the 3 cases. We expanded this experiment by using various numbers of samples in the gallery to test the performance of the proposed method. In this experiment three scenarios were presented, whereby the gallery consist of 2, 3, and 4 Nu samples respectively. In the probe set for each scenario, two samples were used from each of Nu, CB and CW (see Figure 5.5). These results confirm that reducing the number gallery samples by 1 results in lower accuracy in the non-neutral cases by 4% -6%. But the effect in the case Nu is negligible.

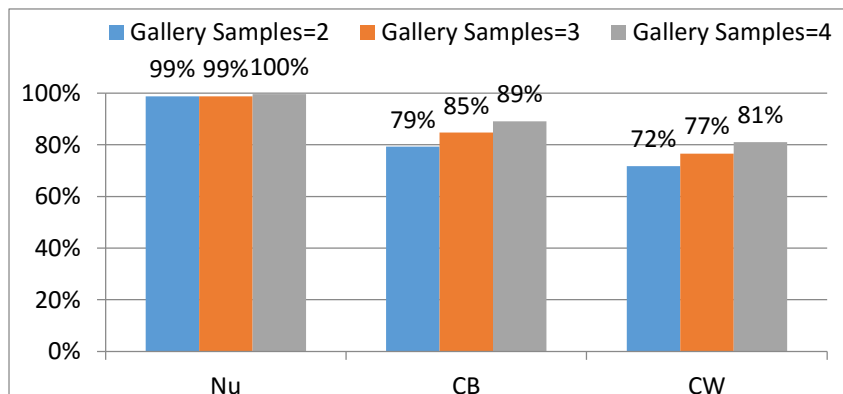


Figure 5-5: Recognition Performance of GSCD-UGR using different number of samples in the gallery (Using CASIA B gait database).

Further examination of these accuracy results, showed that some of the candidates that are misrecognized in the CW case were females but confused as males (see Figure 5.6). Incorporating the gait-based gender classification with the GSCD-UGR scheme is expected to improve accuracy. This will be done over the next two sections.

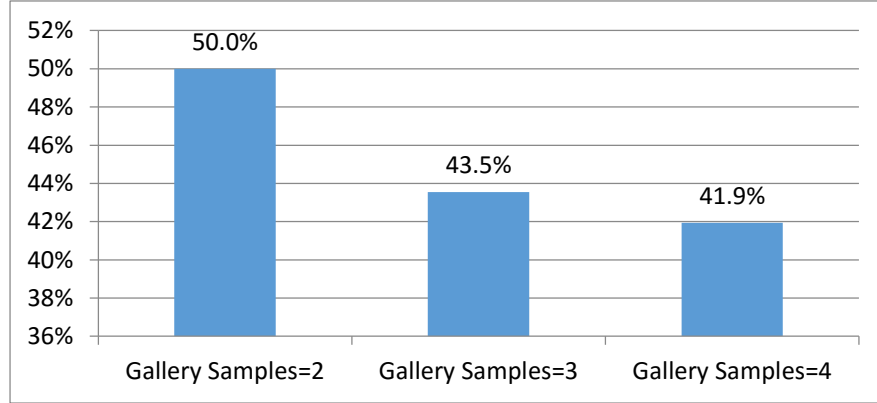


Figure 5-6: The percentage of female candidates that misrecognized and confused as a males in CW gait sequences, using Unrestricted Gait Recognition UGR.

5.4 Unrestricted Gender Classification (UGC)

This method generalize the method developed in chapter for gender classification in neutral gait sequences to include the non-neutral human gait sequence. Again, in all experiments the gallery consist of N_u samples from the CASIA B dataset while the probe consists of 2 samples per person of N_u , CB and CW gait sequences were used as a probe set.

5.4.1 The proposed UGC method

The main difference between UGC method with the scheme developed in chapter 4 is in the feature selection step. The feature extraction is again based on wavelet transformed images. In this stage three types of feature vectors: STM extracted from the LL1 subband, AWM extracted from the LL3 subband , and the DWM extracted from (LH1, HL1 and HH1) non-LL subbands. Note that here we do not use LMD.

After constructing the feature selection, we test the performance of the UGC with each feature vector separately and then we fuse all combination of theses 3 feature vectors. To reduce the dimension of the fused feature vectors and to select the most relevant feature we use the Fisher score as a feature selection method. In the final stage, k-NN

is used as a classification method. Here we present the results for $k=3$, as it gave best results among the 3 values $k=1$, 3, and 5.

Due to imbalanced gender numbers in our experimental CASIA B gait database (93 males and 31 females) we selected a randomly equal subset of 25 males and 25 females of neutral gait sequences for the gallery set. The gallery set was constructed from the first four records of each subjects from the Nu gait sequence only. The remaining records of Nu gait sequences, as well as the CW and CB gait sequences, are used as the probe set for testing. Each experiment was repeated 30 times to cover the entire database, as shown in figure 5.7., and the accuracy rate reported is the average result of the 30 repetitions. The classification result was performed using 10-Fold Cross Validation (FCV).

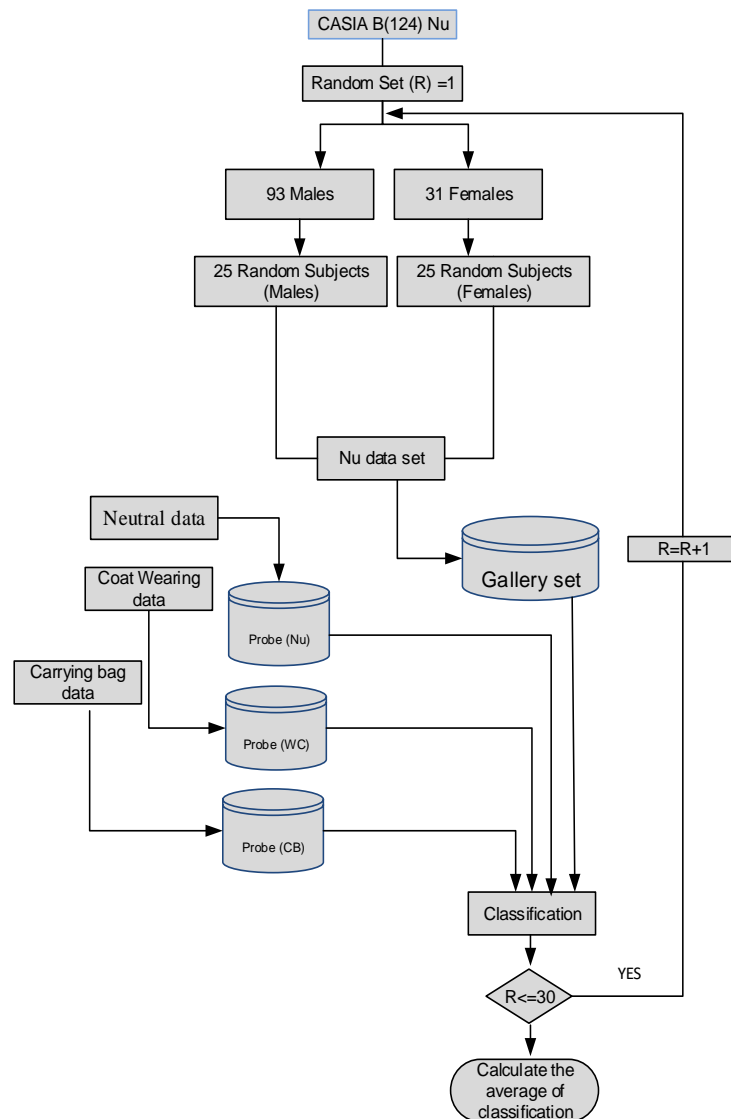


Figure 5-7: Generating gallery and probe set from the CASIA B gait database for gender classification based on Unrestricted Gait Sequences.

5.4.2 Experimental Results

Table 5.2 displays the results of the experiments for the 3 feature vectors, as well as the results of fusing all combinations. These results show that all combinations of feature vectors perform better in the Nu case than in the non-neutral case (with more success in the CB than in the CW case). Among the single feature vectors the STM yields the best gender classifications across all cases. When Fusing 2 feature vectors the accuracy improved, with STM+AWM achieving best gender recognition in the neutral case while the STM+DWM resulted in best achievement in the non-neutral cases. Finally, fusing all feature vectors yielded the best gender classification.

Table 5-2: Recognition Performance of Unrestricted Gender Classification (UGC) for each feature vector (STM, AWM, and DWM) separately and their fusion schemes. (This experiment is based on CASIA B database)

	Nu (%)	CB (%)	CW (%)	Average
STM	94.6	86.9	80.4	87.3
AWM	83.3	68.9	57.2	69.8
DWM	85.3	75.4	58.7	73.1
STM+AWM	97.0	88.5	78.7	88.1
STM+DWM	96.6	89.0	84.0	89.9
AWM+DWM	92.7	82.3	61.5	78.8
ALL	97.7	90.1	87.5	91.8

Figure 5.8, illustrates that the averaging of the results the 30 repetition of the experiments is a good estimate of the genuine accuracy because the standard deviations are very small. Here we only presented the outcome for the fusion of all features.

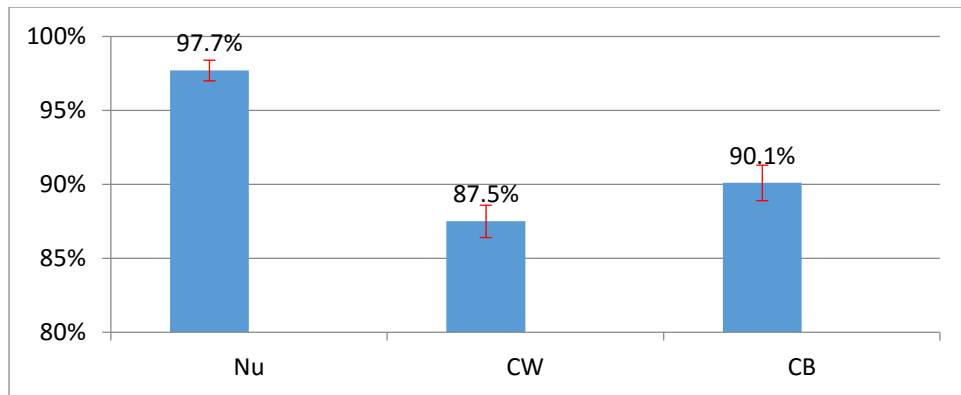


Figure 5-8: Recognition Performance (Average and standard deviation) of Unrestricted Gender Classification using the combination of STM, AWM, and DWM .

Table 5.3, presents a comparison of the performance of our UGC with the only published work that we found in the literature that conduct gender classification in unrestricted case. It can be seen that our scheme outperform the scheme of (Hu, et al., 2010).

Table 5-3: Performance Comparition of UGC with results published in the literature using CASIA B database

Method	CASIA B Dataset	Nu (%)	CW (%)	CB (%)	Average
(Hu, et al., 2010)	31 M & 31 F	96.7	83.8	88.7	89.7
UGC	93 M & 31 F	97.7	87.5	90.1	91.7

5.5 Incorporating Gender classification into GSCD-UGR

In this section we shall incorporate the UGC into the GSCD-UGR in a way that knowledge of the case of walking and gender can provide better chance to improve accuracy of the non-neutral cases. Figure 5.9 illustrates the block diagram of the proposed algorithm. The proposed algorithm is start at GSCD-UGR to obtain the following; (1) Detect the case of the gait sequences (Nu, CB or CW), (2) Recognized Person (RP) and (3) Gender label of RP (G1).If the GSCD detected as CW we do the following:-

- Gender classification based on UGC (G2).
- If G1 is Male and G2 is Female, then we create a threshold (Thr.) based on the same way which proposed in chapter 4 (Section 4.2), and we compare it with the distance of RP (DRP) from the probe sample.
- If DRP is greater than Thr. we apply UGR by using only the Female candidates in the gallery, otherwise we use UGR based on all candidates (Male &Female) in the gallery.

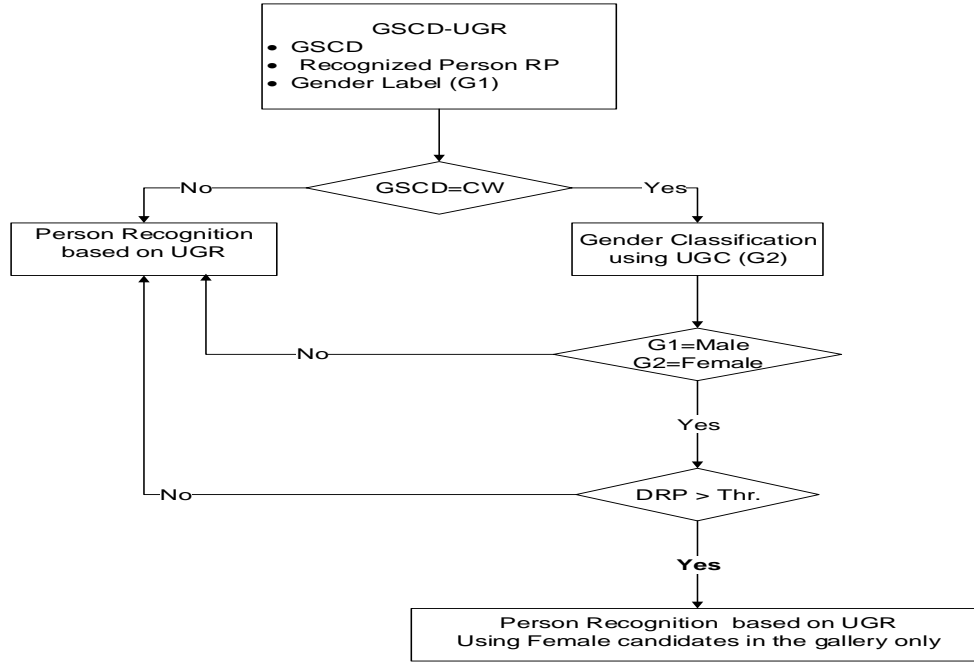


Figure 5-9: Overview model for combining GSCD, UGC, and UGR (DRP is the Distance of Recognized Person from the probe sample)

From this algorithm, UGC trained based the neutral gait sequences only using 25 subjects per 4 records randomly from each of the male and female gait sequences. It is clear that only in the CW case we may have different accuracy rate than the GSCD-UGR algorithm of section 5.3. We therefore present the recognition rate of this new algorithm only for the CW case in Table 5.4. We can see that, the accuracy rates improve by more than 4% when the gender classification was incorporated. After using gender classification with GSCD-UGR the recognition accuracy of Nu, CB and CW are 99.6%, 89.1% and 85.5% respectively and the average is 91.4% (See Figure 5.5).

Table 5-4: Recognition Performance of Unrestricted Gait Recognition (UGR) based on CW, with and without Unrestricted Gender Classification (UGC). Using CASIA B database.

N. Gallery	N. Probe	UGR -CW (%)	(UGR+ UGC) -CW (%)
2	4	71.8	75.8
3	3	76.6	81.1
4	2	81.1	85.5

5.6 Unrestricted Gait Schemes using Gait Energy and Entropy

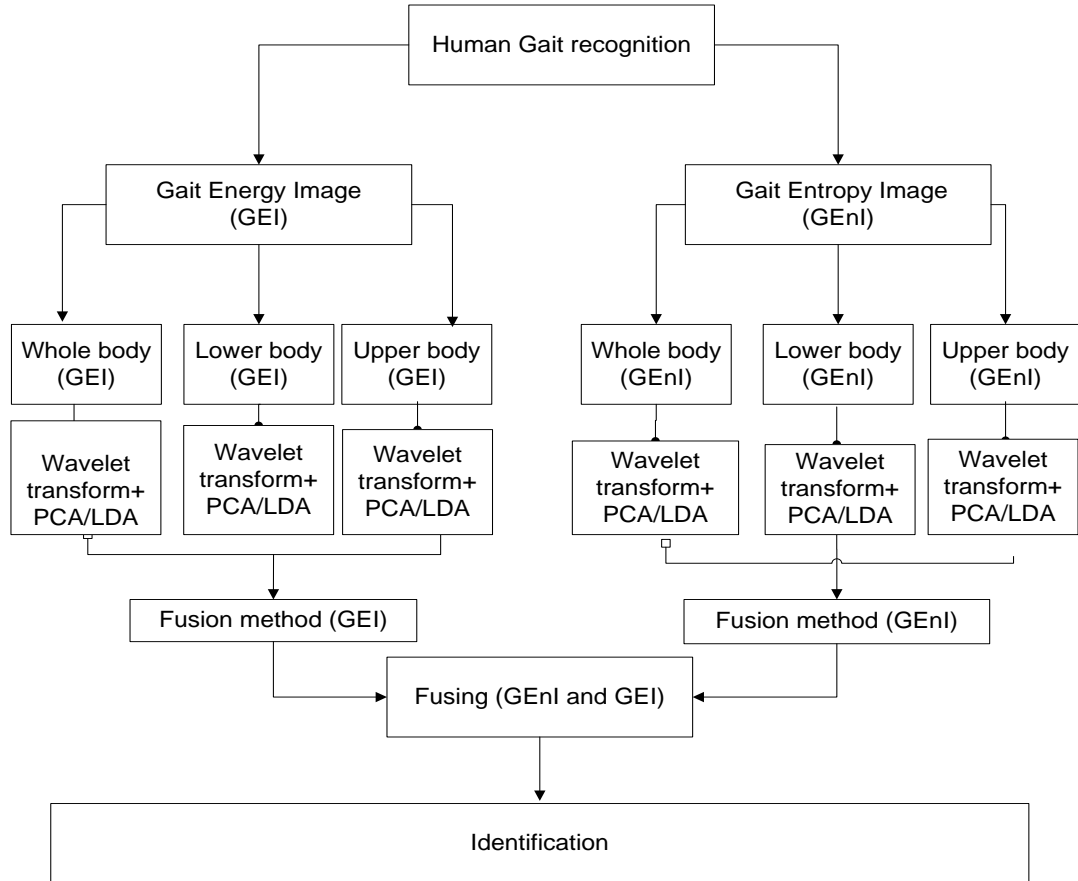
Image (UGS)

Having nearly exhausted the different ways of analysing the wavelet-based spatio-temporal model of features for gait recognition in the neutral as well as unrestricted cases, we now turn our attention to investigate the use of the traditional features of Gait Energy Image (GEI) and the Gait Entropy Image (GEnI) for gait recognition in the unrestricted gait sequences. These two images are obtained from all the frames in a gait cycle video (for detail description of these features, see chapter 3, Section 3.2.4 and Section 3.2.5). These two images have been widely used in recent appearance-based on gait recognition algorithms because of its simplicity and effectiveness (Han & Bhanu, 2006) and (Jeevan, et al., 2013). The limitation of GEI and GEnI is the lacking of robustness to deal with covariate conditions which affect the static areas of human body. Although GEnI is better than GEI in dealing with such problem, but still is not enough to provide good performance especially in CW gait sequences case. As mentioned earlier in this chapter, we assume that the body is divided into 2 parts using the golden ratio proportion (Herman, 2007). Recent studies highlighted the importance of static shape information in the Motion-based gait recognition approaches (Bashir, et al., 2010). Recall that the lower half of body provides gait relevant information and is not significantly affected under covariate condition. However using lower part only will neglect some critical information in the upper half body that may assist gait recognition performance. Using the whole body may also improve the performance in Nu gait sequence and in some cases of CB gait sequences; while the performance of CW goes down by using the whole body and this is presented in our experimental results (see Figure 5.2). The investigations in this section, are therefore work on the GEI and GEnI generated from different parts of the human body. Therefore, we shall investigate the use of combination of these two types of features but extracted from: the lower half, the upper half, and the whole body separately.

5.6.1 UGS-(GEI & GEnI)

Traditionally, each of GEI and GEnI are obtained from the spatial domain of the frames of a gait cycle video. In this section we propose an unrestricted gait recognition using various combinations of GEI and GEnI extracted from parts of the body. In general we have 3 GEI and 3GEnI images Lower GEI (LGEI); Upper GEI (UGEI);

Whole GEI (WGEI); Lower GEnI (LGEI); Upper GEnI (UGEI); and Whole GEnI (WGENI). Moreover, instead of using the spatial domain of these two images, we shall use the wavelet transforms of these images at level 3. In other words, for each part image a feature vector is formed by the concatenation of the 4 level 3 wavelet subbands (LL3, LH3, HL3 and HH3) of the parts GEI/GEnI images. Dimensions reduction, here will be based on PCA. Figure 5.10, illustrate the way these schemes are implemented, followed by the steps of the algorithm.



The algorithm

1. Extract the GEI (and GEnI) from the input gait cycle video, as described in section 3.2., but restricted to 3 body parts (LGEI, UGEI, WGEI; LGENI, UGENI, WGENI)
2. Wavelet transform were applied on each of the above constructed images, and in each case concatenate all the 4 subbands to form the feature vector representing the given part.
3. Reduce the dimension of the feature vectors using PCA.

Figure 5-10: Unrestricted gait Schemes (UGS) using Gait energy and Gait Entropy images.

5.6.2 Description of Experiments

Having defined and extracted the various pairs of Energy/Entropy Image based feature vectors from the various parts, we conducted a gait recognition experiments for each body part represented by the reduced dimension feature vectors, and followed the same protocols before (i.e. first 4-Nu samples per person in Gallery, and 2 samples per persons for the 3cases Nu, CB and CW). Just as before, we also use other protocols by including different gallery scenarios. In these experiments, the NN with cityblock distance function were used as a classification method. The performance of the proposed method has been evaluated based on CASIA B gait database.

5.6.3 Experimental Results

First of All, we present conducted a comparison of using PCA and LDA order to demonstrate that PCA is sufficient for our purpose. However, we restricted the experiment to the whole body feature vectors WGEI and WGenI. The experimental result shown in Figure 5.11, do indeed demonstrate the above observation. Hence, in the rest of the experiments we will only be using PCA for dimensions reduction method.

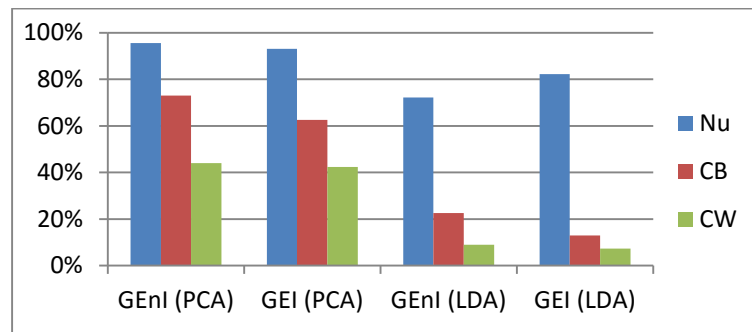


Figure 5-11: Recognition Performance of UGS based on WGEI and WGenI using PCA and LDA separately.

The next set of experiments tested the performance of the various UGR schemes depending the adopted feature vectors which includes the 3 single feature vectors (LBP, UBP WBP) and the score level fusion of any two of them as well as fusion of all. The results are shown in Table 5.5 for both features extracted for GEI and GEnI images. With two exceptions, the GEnI image provides more discriminating features than the GEI image. The only two exceptions are the UBP and WBP+UBP features, in the CW case, where these schemes performed better in the GEI than in GEnI. The

best performing GEnI scheme in the Nu case was WBP (96.4%); in the CB case it is the WBP+LBP+UBP (78.6%); and in the case CW it is the LBP (66.1%).

Table 5-5: Recognition Performance of Unrestricted Gait Schemes (UGS) based on GEI, GEnI, and different combining between them.

		Nu (%)	CB (%)	CW (%)	Average (%)
GEI	WBP	93.2	60.9	42.3	65.5
	UBP	83.9	37.1	15.7	45.6
	LBP	77.8	48.4	58.9	61.7
	WBP+UBP	94.0	58.1	33.1	61.7
	WBP+LBP	91.9	68.6	62.1	74.2
	UBP+LBP	94.4	66.5	52.0	71.0
	WBP+UBP+LBP	95.2	67.7	53.2	72.0
GEnI	WBP	96.4	70.6	46.8	71.2
	UBP	87.9	48.0	14.5	50.1
	LBP	83.9	52.8	66.1	67.6
	WBP+UBP	96.0	71.0	32.7	66.5
	WBP+LBP	95.6	77.0	64.9	79.2
	UBP+LBP	95.2	71.8	56.1	74.3
	WBP+UBP+LBP	95.6	78.6	55.2	76.5

Note that results achieved by best combination of features in GEnI are comparable the State of the art in each of the cases, but are significantly outperformed by the results achieved in the UGR (section 5.5) in each of the walking cases. A full comparison with the state of the art will be conducted at the end of this chapter. However, we conducted another set of experiments to fuse, again at the score level, the various combinations of feature vectors extracted from both the GEI and GEnI to see if better results can be achieved. The results from fusing the 7 different feature vectors in GEI with their corresponding feature in the GEnI are shown in Table 5.6. We also conducted an extra “ensemble” pattern recognition scheme where identification decision is based on the simple majority rule applied to these 7 fused schemes. Here, simple majority rule is based on taking the decision that agreed by the largest number of the classifiers, but if all identified a different person the system takes the decision associated with the lowest score. The results shows that accuracy has been improved more, especially by fusing (WGEI + LGEI) with (WGenI + LGenI) achieving 83.2% as an average of the three walking cases. The amount of improvement is nearly 4%

for the CB case, and 5% for CW case, but 0% in the Nu case. The ensemble scheme, has not improved the accuracy but had achieved the same accuracy as the fused WBP+LBP+UBP for the Nu case only. But these results, including the ensemble scheme, are still way below what was achieved in section 5.4.

Table 5-6: Recognition Performance of UGS based on fusing GEI and GEnI with each of WBP, LBP, and UBP and different combinations between them.

GEI+GEnI	Nu (%)	CB (%)	CW (%)	Average (%)
WBP	95.6	77.0	46.4	73.0
UBP	88.3	50.0	18.6	52.3
LBP	84.3	61.3	69.8	71.8
(WBP+UBP)	96.0	73.4	35.5	68.3
(WBP+LBP)	96.0	82.3	71.4	83.2
(UBP+LBP)	95.6	75.0	59.3	76.6
(WBP+LBP+UBP)	96.4	80.7	58.9	78.7
Ensemble method	96.4	81.5	64.5	80.8

We conducted more experiments to consider the effect of including different number of samples (not necessarily all Nu) in the gallery and probe set. This experimental applied on UGS-(GEI &GEnI) using (WBP+LBP) which provided the best performance among other fused feature vectors in the standard scenario. The experimental result confirms our expectation that having the same gait sequences in the gallery and probe set will increase the chance of recognition (see Table 5.7).

Table 5-7: Recognition Performance of UGS based on combination of GEI and GEnI (WBP+LBP) with different gait sequences in the gallery and probe set.

GEI & GEnI (P-G)	Nu-Nu (%)	CB-Nu (%)	CW-Nu (%)	CW-CW (%)	CB-CB (%)	CB-CW (%)	Average (%)
Gallery(G) No.	4(Nu)	4(Nu)	4(Nu)	1(CW)	1(CB)	2 CB	
Probe(P) No.	2(NU)	2(CB)	2(CW)	1(CW)	1(CB)	2 CW	
(WBP+LBP)	96.0	82.3	71.4	86.3	79.03	32.26	74.5

5.7 Combining UGS-(GEI & GEnI) with UGR

After testing the performance of the two different unrestricted gait recognition methods in section (5.5 and 5.6), we considered fusion of the two methods (for some feature vectors) at the score level. The tested combination are based on fusing the best performing features in UGS-(GEI&GEnI) using (WBP+LBP) with 3 different UGR schemes discussed in section 5.1-5.5 (S2-UGR, GSCD-UGR, and GSCD-UGR

+UGC). The results of these experiments are shown in Table 5.8. These results show that the fusion has not improved the unrestricted gait recognition at all above what was already achieved by the (GSCD-UGR +UGC) scheme.

Table 5-8: Recognition Performance of gait recognition based on fusing UGS and UGR using score level fusion).

	Nu (%)	CB (%)	CW (%)	Average (%)
UGS-(GEI & GEnI) + S2-UGR	98.0	87.1	80.2	88.4
UGS-(GEI & GEnI) +(GSCD-UGR)	99.6	89.1	80.2	89.6
UGS-(GEI & GEnI) +(GSCD-UGR +UGC)	99.6	89.1	84.3	91.0

We then fused the two scheme UGS-(GEI & GEnI) with (GSCD-UGR +UGC) at the decision level, i.e. by comparing the normalised distances between the probe sample and the nearest gallery template calculated by the two methods and choosing the identity of the subject of the minimum distance. The experimental results, Table 5.9, show that decision level fusion outperform the score fusion in all cases for the non-neutral cases except in the case of the UGS-(GEI & GEnI) OR S2-UGR scheme. For the CB case the decision-based fusion resulted in improved accuracy by nearly 3% when GSCD were used, and for the CW case the decision based fusion improved by about 1.7%. In conclusion, we should recommend the decision level fused UGS-(GEI & GEnI) OR (GSCD-UGR +UGC) scheme.

Table 5-9: Recognition Performance of gait recognition based on fusing UGS and UGR using decision level fusion.

	Nu (%)	CB (%)	CW (%)	Average (%)
UGS-(GEI & GEnI) <u>OR</u> S2-UGR	98.0	84.3	81.9	88.0
UGS-(GEI & GEnI) <u>OR</u> (GSCD-UGR)	99.6	91.9	81.9	91.1
UGS-(GEI & GEnI) <u>OR</u> (GSCD-UGR +UGC)	99.6	91.9	86.3	92.6

5.8 Comparison of our methods with the State Of The Art

In this section we compare our results with a set of existing unrestricted gait recognition for both type of protocols (gallery consisting of only Nu samples for, and gallery containing non-neutral samples). Table 5-10 (Table 5.11) displays the results of the first (second protocol for all our single and fused schemes and the results from 7 existing unrestricted gait identification schemes published over the period 2009-2014.

Table 5-10: Performance Comparison between our methods and the State of the Art – standard protocol. (Based on CASIA B gait database (124 subjects), for each subject, the first 4 Nu gait sequences used in the gallery and 2 Nu, 2 CB, and 2 CW gait sequences used as a probe). (All figures are percentages).

Probe- Gallery	Nu-Nu (%)	CB- Nu (%)	CW- Nu (%)	Average (%)
(Jeevan, et al., 2013)	93.4	56.1	22.4	57.3
(Yogarajah, et al., 2011)	98.4	93.4	44.0	78.6
(Bashir, et al., 2009) GEI	99.4	60.2	30.0	63.2
(Bashir, et al., 2009) GEnI	98.3	80.1	33.5	70.6
(Bashir, et al., 2010)	100	78.3	44.0	74.1
(Sivapalan, et al., 2012)	100	68.5	80.3	82.9
(Hsia, et al., 2013)	99.0	75.0	73.0	82.3
UGS-(GEI & GEnI)	96.0	82.3	71.4	83.2
S2-UGR	99.2	82.7	81.1	87.6
UGS-(GEI & GEnI) + S2-UGR	98.0	87.1	80.2	88.4
GSCD-UGR	99.6	89.1	81.1	89.9
GSCD-UGR +UGC	99.6	89.1	85.5	91.4
UGS-(GEI & GEnI) +(GSCD-UGR)	99.6	89.1	80.2	89.6
UGS-(GEI & GEnI) +(GSCD-UGR +UGC)	99.6	89.1	84.3	91.0
UGS-(GEI & GEnI) <u>OR</u> (GSCD-UGR +UGC)	99.6	91.9	86.3	92.6

Table 5-11: Performance Comparison between our methods and the State of the Art – mixed protocol.

Probe –Gallery	Nu-Nu (%)	CB-CB (%)	CW-CW (%)	CB-Nu (%)	CW-Nu (%)	CB-CW (%)	Average (%)
(Hu, et al., 2013) LF-AVG	71.4	63.1	60.7	13.1	20.2	11.9	40.1
(Hu, et al., 2013) LF-oHMM	63.8	31.8	21.4	19.7	22.6	9.1	28.1
(Hu, et al., 2013) LF-iHMM	94.0	64.2	57.1	45.2	42.9	22.6	54.3
(Jeevan, et al., 2013) GPPE	93.4	62.2	55.1	56.1	22.4	17.9	51.2
(Jeevan, et al., 2013) ShGEnI	92.3	65.3	55.1	56.1	26.5	18.9	52.4
(Kusakunniran, 2014)	95.4	73.0	70.6	60.9	52.0	29.8	63.6
UGS-(GEI & GEnI)	96.0	79.0	86.3	82.3	71.4	32.3	74.5
S2-UGR	99.2	54.8	62.1	82.7	81.1	65.7	74.3

These two tables demonstrate beyond any doubts that all our fused methods (at the score or at the decision level) outperform all existing schemes. Moreover the result that provided using UGS-(GEI & GEnI) OR (GSCD-UGR +UGC), 92.6% is significantly better than 82.9% that provided in (Sivapalan, et al., 2012), with $p = 5 \times 10^{-6}$.

5.9 Conclusions

In this chapter we investigated gait recognition in the unrestricted scenario which includes walking in non-neutral circumstances (CB and CW) in addition to the neutral Nu case. When attempted to extend the use of the NGR scheme, developed in the previous chapter, for gait recognition in the unrestricted case (i.e. UGR) the accuracy for the CB and CW were way below the accuracy for the Nu condition. We extended the sets of features used in the NGR, which were entirely in the LL3 wavelet subbands, by sets of features extracted from all high frequency subbands after wavelet decomposition to level 3. Two versions of the extended features were tested: S1 was based on whole body and S2 was based on the lower body part only. The accuracy for the S1-based UGR improved significantly for the CB case but only marginally for the CW, while the accuracy of S2-based UGR, improved by large margins for both CB and CW cases. But the S2-based UGR in Nu and CB were slightly lower than the S1-based scheme. This has motivated the development of the gait sequence case detection (GSCD) procedure. Incorporating the GSCD into the UGR has led to improved accuracy but not by much in the CW case. Having found that in the CW case many errors male and females were confused. By further incorporating the Gender classification scheme, developed earlier have significantly improved the accuracy of UGR. We further investigated the use of different sets of features extracted from the commonly used GEI and GENI compactification images of the gait sequences captured from different part of the body. Again these sets of features were wavelet based rather than spatial domain features that is adopted by other researchers. Recognition based on schemes that use these sets of features, or their fusions at the score level, where all termed as UGS schemes. UGS schemes that were based on fusing features from GEI and GENI resulted in high accuracy but not as high as those achieved by the UGR schemes. Fusing the UGR schemes with their equivalent UGS schemes at the feature level did not improve on the UGR schemes, but when we used decision based fusion to combine them we achieved the highest accuracy in comparison with the state of the art. We can attribute this success to the fact that for the non-neutral cases were recognised using fusion of features that were extracted from lower body part. In the next chapter we will investigate gait recognition using Kinect sensor, where most of the features are provided from skeleton points

determined by the camera. This could reveal the effect of errors in the automatic pre-processing, that we used so far, on the accuracy of gait recognition.

Chapter Six

Gait recognition using Kinect sensor

In the previous chapters gait recognition methods were proposed and tested using the CASIA B database where the video data were captured by a conventional camera. The availability of the new Kinect sensor for video data gathering, originally designed for game applications enables new research opportunity for many purposes like gait recognition for real-time security surveillance, for monitoring elderly people, as well as estimating progress in recovery from injuries that impair walking style. Among the investigations made in this chapter, is the use of the Kinect sensor for gait recognition with Windows SDK in which provides high quality tracking information. Here we used the Kinect sensor to obtain a 3D human skeleton on which 20 joint points are output to be used for gait recognition from which highly accurate distance features similar to those in our STM feature set. As we have seen when developing our automatic procedures, some features are difficult to extract with high accuracy from normal sequence videos. In this chapter we propose gait recognition based on the Kinect sensor. The proposed method is divided into two parts: the first uses the Kinect sensor for recognising neutral gait sequences and the second deals with using the Kinect sensor for recognising human on gait sequences under non-neutral circumstances.

6.1 Human Skeleton Tracker

Kinect technology is a human machine interface that has been attracting growing interest from researchers. It is simple and easy to use and it holds three crucial bits that work together to detect human motion and raise physical image from human body on the screen: a RGB color VGA video camera, a depth sensor, and a multi-array microphone. Red, green, and blue color components and also body-type and facial features are detects by the camera. Kinect has a frame rate of 30 fps.

The depth sensor holds a monochrome CMOS sensor and infrared projector that assistance make the 3D imagery throughout the room. It likewise measures the separation of each perspective of the human body by transmitting invisible near-

infrared light and measuring its "time about flight" then afterward it reflects off the object (Zhang, 2012).

In this chapter we propose and test the performance of a gait recognition system based on using the Kinect sensor to collect automatic tracking information from video recordings of persons walking on a side view of the camera. Currently there is no standard and useful application for recording the human skeleton by Kinect, and consequently we had to create an application to get skeleton points, to track 20 joint points of the human skeleton. This application automatically saves these points in an Excel file. The 20 skeletal points are spread over the human body all the way from the head through the hips region down to the feet (see table 6.1 for precise description).

Table 6-1: Joint points name and numbers from Kinect Sensor (N is number).

Joint	Joint Name	Joint	Joint Name	Joint N.	Joint
1	Hip Centre	8	Hand Left	15	Ankle Left
2	Spine	9	Shoulder	16	Foot Left
3	Shoulder	10	Elbow Right	17	Hip Right
4	Head	11	Wrist Right	18	Knee
5	Shoulder Left	12	Hand Right	19	Ankle
6	Elbow Left	13	Hip Left	20	Foot Right
7	Wrist Left	14	Knee Left		

For each video frame, Kinect outputs the coordinates of these 20 points which will be recoded in an excel file. We used the application to collect a new databases to be used in evaluating the proposed gait recognition.

Before we began creating the intended databases, we tested the accuracy of the Kinect's distance measurements between the Kinect and a subject. To do that we recorded 10 short videos for 5 subjects walking at different distances (3.5m, 3m and 2.5m). Then, we compared the accuracy of the Kinect distances, and we found that the optimum distance between the Kinect and the subject, in terms of accuracy of the extracted distances, is 2.5m.

6.2 Kinect Database

Kinect sensor is relatively new, to the best of our knowledge, currently there is no relevant public dataset for gait recognition that was made with Microsoft Kinect and provide skeletal information. Although in (Borràs, et al., 2012) the data captured using a Kinect sensor and presented as a public database, however there was no attempt at inferring or using skeletal information and only depth data and regular video were provided. Moreover the carrying bags (on their back, or over shoulder), or wearing coats (long and short) challenges that addressed in this thesis and labeled as non-neutral has not addressed in this database. Therefore it was necessary to build our own database using the Kinect camera. We recorded two different datasets that collected in slightly different locations (See figure 6.1); the first dataset was created based on neutral gait sequences, and this will be referred to thereafter as the Kinect database-1. In this databased we video recordings for 20 participants (16 males and 4 females) walking in front of the Kinect sensor from right to left at an angle of 90 degrees. The Kinect sensor was placed on a table at a height of 0.6m, the angle of the Kinect with no inclination. The Kinect sensor records approximately 30 frames per second. The participants were asked to walk normally 10 times, providing a total of 200 recordings.



Figure 6-1: Kinect database locations.

The second dataset, named Kinect database-2, recorded videos for another set of 20 participants, in a different location, but also the purpose was to include recording of persons walking in neutral, as well as non-neutral cases. For each subject, there are 25 walking sequences, consisting of 5 Neutral (i.e. Set Nu), 5 Wearing Long Coats (i.e. Set WLC), 5 Wearing Short Coats (i.e. Set WSC), 5 Carrying Bag over Shoulder (i.e. Set CBS) and 5 Carrying Bag on Back (i.e. Set CBB). This means that database-2 includes a total of 500 videos. Again, in this dataset the participants were asked to

walk in front of the Kinect sensor, which was placed a table at a height of 0.6m with no inclination, from right to left at an angle of 90 degrees.

6.3 Gait Recognition Using Kinect Sensor

Here we propose a method for gait recognition based on the 20 skeletal points model provided by the Microsoft Kinect sensor. As we explained above, a Kinect sensor with Kinect for Windows SDK v 1.6 provides a high quality human skeleton for up to two persons. In general, our system consists of three stages, as depicted in figure 6.2. In the first, pre-process the videos to extract the 20- skeletal point for each of the frames constituting a single gait cycle using local maxima of the bounding box width sequence. In the second stage, we compute two sets of distance features: Horizontal Distance Features (HDF) and Vertical Distance Features (VDF). Finally, the use of classification method(s) is the final stage. This performance of this model on both recorded databases, but in the first set of experiments only neutral gaits recognition is tested. Moreover, in the neutral gait experiments we shall test the performance of our scheme with two classifiers, the Linear Discriminant Classifier (LDC) as well as the Naive Bayes classifier in order to choose the best one for unrestricted gait recognition.

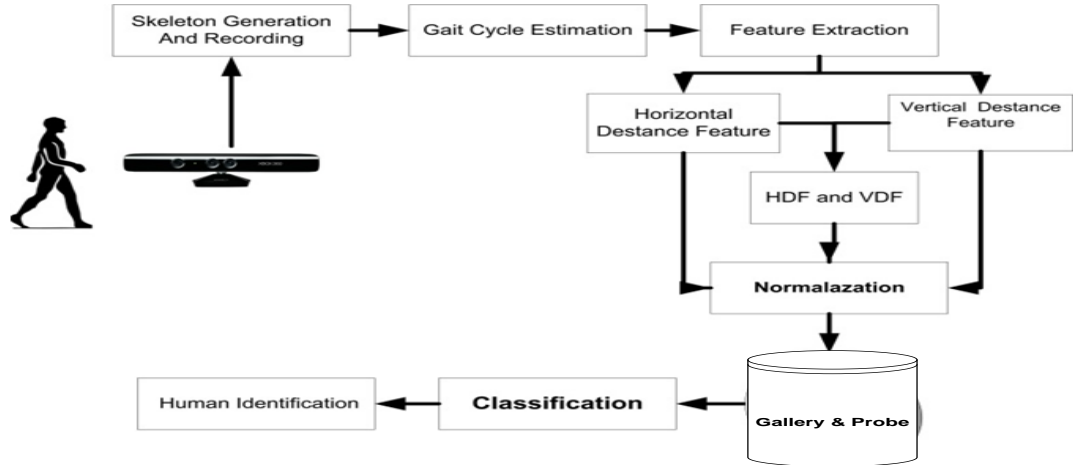


Figure 6-2: Overview of gait recognition system using Kinect Sensor

6.3.1 Feature extraction based on Kinect

The proposed feature sets, used in previous chapters, were extracted from human silhouettes that were generated based on conventional camera. In this study we extract

the features based on Kinect sensor, thus the process of feature extracting is totally different from the techniques were used normal cameras. As already mentioned, the Kinect sensor provides 20 joint points located on the skeleton of the walking participant, and output in the form of the pair of x -coordinate and y -coordinates for the 20 joints. These points are given an index from 1 to 20 as depicted in Figure 6.3. Using the xy -coordinates of these points, we compute two sets of dynamic features; the first set is the horizontal distance feature and the second set is the vertical distance feature.

6.3.1.1 Horizontal Distance Feature (HDF)

The Kinect based feature HDF feature vector is based on measuring the changes in distance between the skeleton joints, in the x -direction across the frames of a single gait cycle. This feature vector is defined in terms of the distributions of four features set; the step length (HD1), the distance between the right and left knees (HD2), the distance between the right and left wrists (DH3) and the distance between the right and left shoulders (HD4) (see Figure 6.3). Here, instead of measuring the distance between the left and right feet to detect step length, we used the left and right ankles for better accuracy. Then, we calculate the mean, standard deviation, and skewness for each of these measurements over all frames in one gait cycle. The following set of equations define the 12-dimensional HDF feature vector.

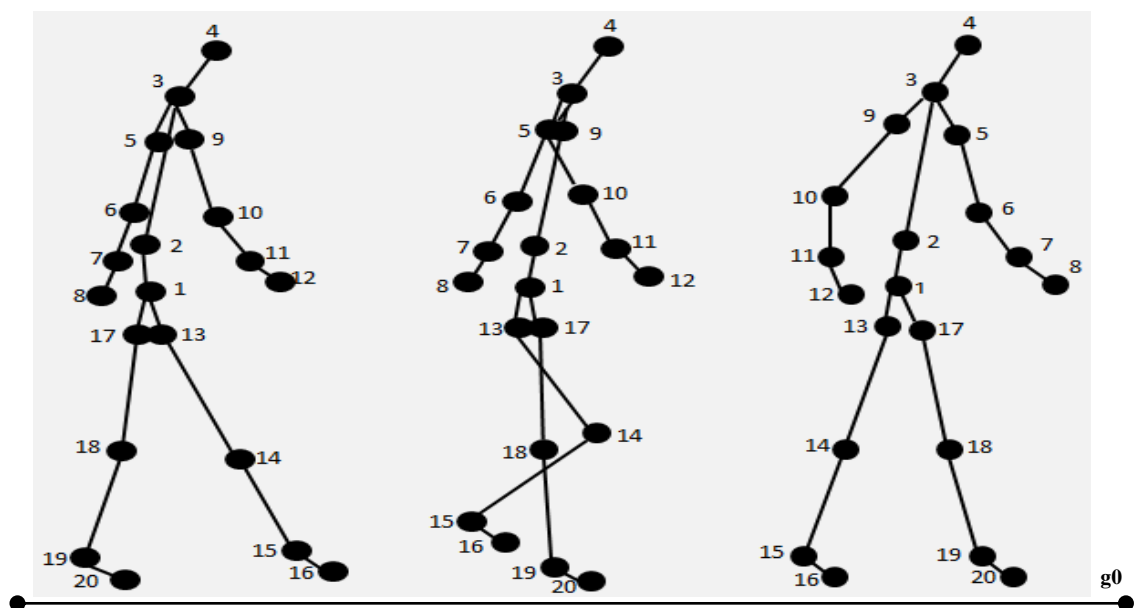


Figure 6-3: Overview of the skeleton points from Kinect Sensor.

$$HD1 = |x_{19} - x_{15}| \quad (6-1)$$

$$HD2 = |x_{18} - x_{14}| \quad (6-2)$$

$$HD3 = |x_{11} - x_7| \quad (6-3)$$

$$HD4 = |x_9 - x_5| \quad (6-4)$$

$$\text{MeanHD} = \{\text{mean}(HD1_{i=1:n}), \text{mean}(HD2_{i=1:n}), \text{mean}(HD3_{i=1:n}), \text{mean}(HD4_{i=1:n})\}$$

$$\text{StdHD} = \{\text{std}(HD1_{i=1:n}), \text{std}(HD2_{i=1:n}), \text{std}(HD3_{i=1:n}), \text{std}(HD4_{i=1:n})\}$$

$$\text{SkewHD} = \{\text{skew}(HD1_{i=1:n}), \text{skew}(HD2_{i=1:n}), \text{skew}(HD3_{i=1:n}), \text{skew}(HD4_{i=1:n})\}$$

$$\text{HDF} = [\text{MeanHD}, \text{StdHD}, \text{SkewHD}] \quad (6-5)$$

Where n is the number of frames in one gait cycle.

6.3.1.2 Vertical Distance Feature (VDF)

VDF is the other feature vector extracted from the coordinates of the 20 joints captured by the Kinect sensor, this feature is constructed from the changes in distance between certain skeleton joints measured along the vertical y-coordinate. Here, we proposed six distances to represent this feature vector, as follows: the participant's height (VD1), the height of the right wrist (VD2), the height of the right shoulder (VD3) and the height of the right and left ankles (VD4 and VD5). Finally, we created a triangle based on the centre of the hips and the distance between the right and left feet (VD6). Again, we calculate the mean, the standard deviation, and the skewness of all of these measurements in one gait cycle. The following set of equations define the 18-dimension VDF feature vector (see Figure 6.3).

$$VD1 = |y_4 - g_0| \quad (6-6)$$

$$VD2 = |y_{11} - g_0| \quad (6-7)$$

$$VD3 = |y_9 - g_0| \quad (6-8)$$

$$VD4 = |y_{19} - g_0| \quad (6-9)$$

$$VD5 = |y_{15} - g_0| \quad (6-10)$$

$$VD6 = 1/2 |x_{20} - x_{16}| * y_1 \quad (6-11)$$

$$\text{MeanVD} = \{\text{mean}(VD1_{i=1:n}), \text{mean}(VD2_{i=1:n}), \text{mean}(VD3_{i=1:n}), \text{mean}(VD4_{i=1:n}), \text{mean}(VD5_{i=1:n}), \text{mean}(VD6_{i=1:n})\}$$

$$\text{StdVD} = \{\text{std}(VD1_{i=1:n}), \text{std}(VD2_{i=1:n}), \text{std}(VD3_{i=1:n}), \text{std}(VD4_{i=1:n}), \text{std}(VD5_{i=1:n}), \text{std}(VD6_{i=1:n})\}$$

$$\text{SkewVD} = \{\text{skew}(VD1_{i=1:n}), \text{skew}(VD2_{i=1:n}), \text{skew}(VD3_{i=1:n}), \text{skew}(VD4_{i=1:n}), \text{skew}(VD5_{i=1:n}), \text{skew}(VD6_{i=1:n})\}$$

$$\text{VDF} = [\text{MeanVD}, \text{StdVD}, \text{SkewVD}] \quad (6-12)$$

Where n is the number of frames in one gait cycle.

6.3.2 Performance of the Kinect Neutral Gait Recognition (KNGR)

The KNGR scheme is based on using the 30-dimensional feature vector (HDF, VDF) digital representation of a person determined from 20 skeletal joint points obtained in single gait sequence of frames captured by the Kinect sensor while the participant is walking under the neutral condition. To test the performance of this KNGR scheme we used the Kinect database-1 which was recorded for this purpose. This experiment deals with recognizing neutral gait sequences based on the Kinect sensor. There were 20 participant in this database, and their recorded raw data includes a set of (x, y) coordinates representing the 20 joint points, for all frames in a gait cycle, captured by the Kinect sensor as described above. In this experiment, each sample is represented by the (HDF, VDF) feature vectors extracted from the raw data record in the database. In total we have 10 samples per person. We followed the tradition in most publications and conducted a 10-fold Cross Validation experiments, i.e. each time we select 1 sample per person for testing and 9 samples per person in gallery. This was repeated 10 times, and each time different test samples are selected. In this experiment we tested the performance of the KNGR using the LDC and Naive Bayes classifiers. In order to determine the contribution of each of the two sets HDF and VDF we tested our system by using each of the 2 components separately and then combined in the feature level to represent one feature vector. The results of this experiment is presented in Figure 6.4. The results show that HDF alone achieve lower recognition rate (57% for the Naive Bayes and 59% for the LDC) whilst VDF provided 84% and 88% recognition accuracy using the Naive Bayes and LDC respectively. However combining HDF and VDF into one feature vector the recognition rate increased to 89% using Naive Bayes

and 93% using LDC. The fact that VDF has better discriminating power than the HDF can possibly be attributed to the fact that walking style is effected more by the amount of effort the person put into raising the legs and hands.

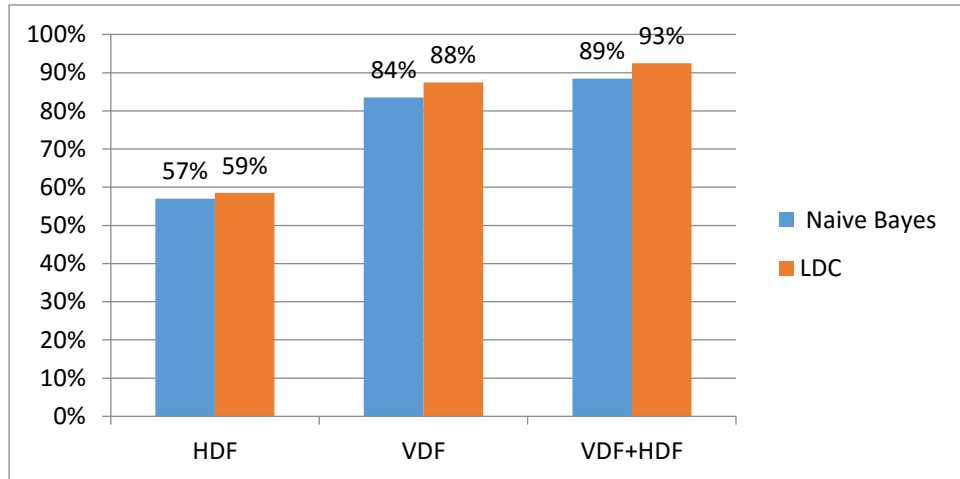


Figure 6-4: Recognition Performance of Kinect Neutral Gait Recognition (KNGR) using Kinect database1 (20 participants, 10 records each).

When we repeated this experiment 20 times by randomly choosing gallery to probe samples according to a 50-50 protocol, the accuracy went down only slightly to (55.3% for HDF, 85.2% for VDF and 90.5% for the combined features) using the LDC classifier. These results justify the dropping of the Naïve Bayes classifiers in the next set of experiments for unrestricted gait recognition. Comparing our approach with those of other researchers is problematic because there is a lack of a benchmark database, due to the unavailability of a similar database in the public domain. Researcher who used a Kinect sensor, including us, created their own database; there are several differences between these databases in terms of the number of participants, the number of records per subject, the distance between subject and sensor, the height of the sensor, and so on. We expect that each of these measurements has an effect on the recognition rate. The use of various classification techniques is another factor making comparisons difficult. Nevertheless we present, in Table 6-2 a comparison of performance between our scheme and a number of methods we found in the literature where the Kinect sensors was used. In the table we give the number of participants in these studies and the achieved accuracy rates.

Table 6-2: Performance Comparison according to recognition rate between proposed method and other methods in the literature. (M is Method and No. is Number)

Related work	Classification M.	No. of Candidates	Recognition
(Ball, et al., 2012)	K-means	4	43.6%
(Sinha, et al., 2013)	ANN	10	90%
(Preis, et al., 2012)	Naive Bayes	9	91%
Our Method	LDC	20	93%

Our approach seems to be practical and has the following positive points. Firstly, we used our database that includes skeleton information for 20 people. Secondly, we proposed two sets of new and meaningful features for human recognition.

6.3.3 Unrestricted Gait Recognition Using Kinect Sensor (KUGR)

In this section, we propose and test the performance of a gait recognition scheme that is used to recognise a person from an unrestricted gait sequence. This method is also based on (HDF, VDF) explained above. Recall that the unrestricted gait sequences under different conditions (Nu, CW and CB). As explained earlier a second database was recorded using the Kinect sensor (Kinect database-2), to be used as our experimental database. Unlike the CASIA-B database, in this database we expanded the number of unrestricted gait cases to distinguish between two cases of carrying bags (carrying on the back or on the shoulder) and two types of coats (long and short).

Performance testing Experiments for database-2

We now extend our Kinect-based gait recognition to the unrestricted walking style. The Kinect sensor provided opportunities to collect gait data for a variety of walking style extending the two non-neutral case into 4 cases to enable the distinction between carrying a bag on the shoulder or on the back, and between wearing a long or a short coat. The new cases include: neutral case (Nu), carrying bag by back (CBB), and carrying bag by shoulder (CBS), wearing short coat (WSC) and wearing long coat (WLC) (see Figure 6.5). Moreover, it also enabled the creation of database-2 which includes more samples for all the 5 walking style, which allow experiments without the limitation of unbalanced class samples as was the case in the CASIA-B.

In the experiments conducted here, we test the performance of KUGR on the Kinect database-2 data, but only using the combined 30-dimensional feature vector (VDF, HDF), and we adopt the LDC classifier. This decision is based on the fact that the

(Feature vector, Classifier) combination resulted in the best performance in the previous section. The reported in experiment is conducted under different scenarios each stipulating the number and cases of samples of gait sequences included in the gallery and probe sets. Table 6.3 specify 5 scenarios for our experiments, where G stands for Gallery set and P stands for Probe set.

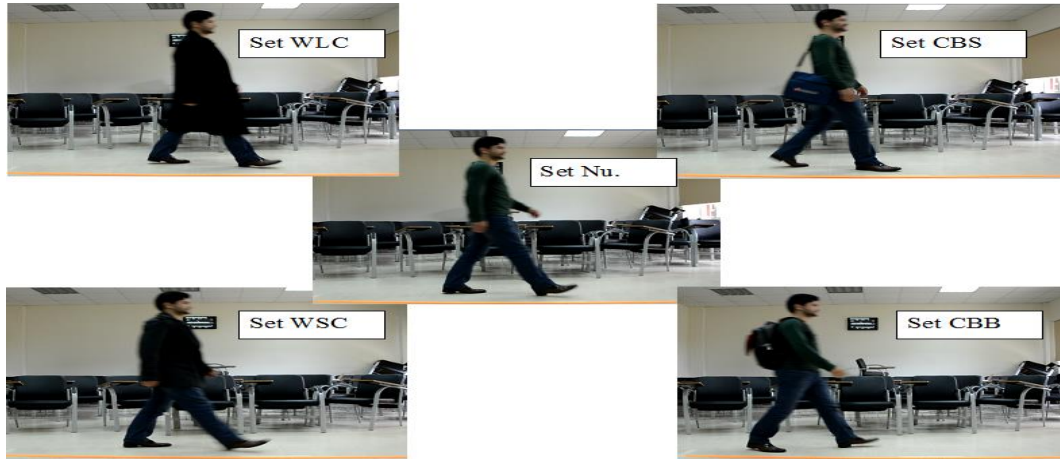


Figure 6-5: Kinect database-2 example.

Table 6-3: Gallery and probe samples set in five scenarios. (G is gallery set and P is probe set).

Scenarios Cases	1		2		3		4		5	
	G	P	G	P	G	P	G	P	G	P
<i>Nu</i>	5	Non	5	Non	5	Non	2	3	1	4
<i>CBB</i>	No	5	1	4	2	3	2	3	1	4
<i>CBS</i>	No	5	1	4	2	3	2	3	1	4
<i>WSC</i>	No	5	1	4	2	3	2	3	1	4
<i>WLC</i>	No	5	1	4	2	3	2	3	1	4
<i>NS</i>	5		9		13		10		5	

The results of these experiments are shown in figures 6.6-6.10 for scenarios 1 – 5, respectively. In Figure 6.6, since the gallery in this scenario is made up entirely of Nu cases, the results represent the percent of accurate recognition for the 20 persons in each of the other 4 non-neutral cases. In each of the other 4 scenarios, the experiment is repeated for different groups of samples per person per case, and the results in the

corresponding figures present the mean of accuracy over all repetitions for each walking case together with the standard deviation of these repetitions.

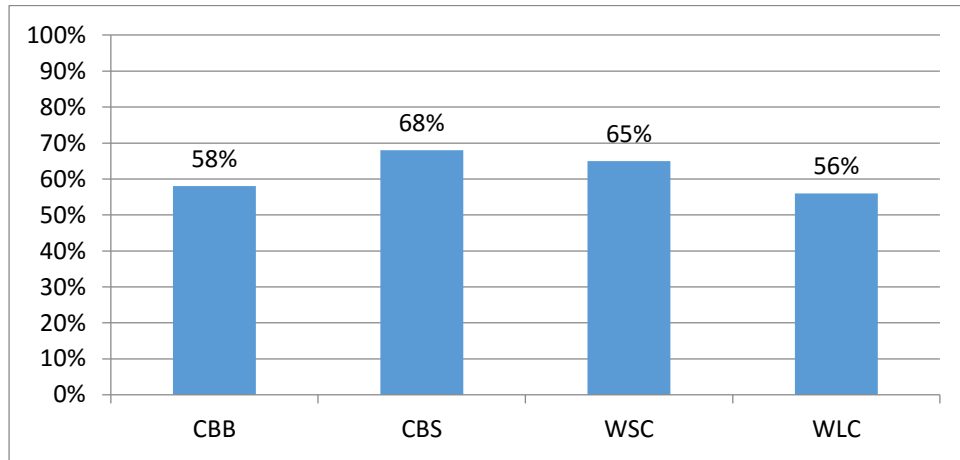


Figure 6-6: Recognition Performance of Unrestricted Gait Recognition using Kinect Sensor (KUGR) based on the first scenario, using Kinect database 2, (20 participants, 25 records each).

The results for the non-neutral cases in this scenario are well below what was achieved for the Nu case. But this is expected because the samples in the gallery are entirely Nu samples. In comparison to what was achieved in Chapter 5, when we tested the performance of the NGR on the non-neutral cases, the performance of the current scheme for the two carrying bag cases are below that of the NGR for CB while the current scheme achieves significantly better results in recognizing the coat wearingcases than the NGR performance for CW (25%). This comparison between the Kinect base scheme and the NGR most likely reflects the differences in the features used but it may also have been affected by the larger number of testing samples.

In the second scenario, where the gallery is supplemented with 1 sample for each of the non-neutral cases, the results show that the recognition result have improved significantly compared to what is achieved in scenario 1 for all the non-neutral cases. Moreover, these results are almost comparable to the outcome of most of experimental testing of the UGR for CW case and only marginally lower than those achieved by UGR for the CB case. The relatively small values of the standard deviations indicates that the mean values are good representative of the accuracy results and also there possibilities of improving accuracy by carefully selecting gallery samples (see Figure 6.7).

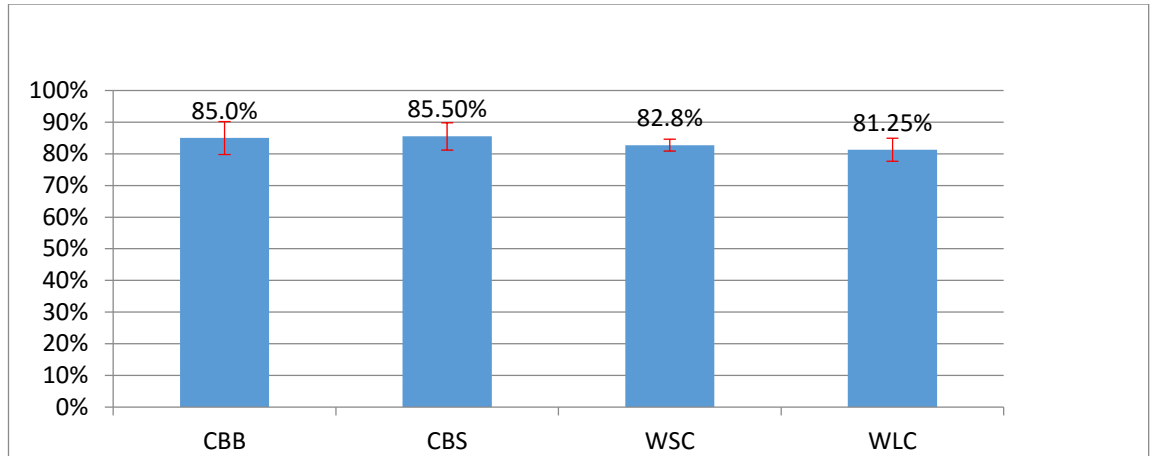


Figure 6-7: Recognition Performance (Average and standard deviation) of KUGR based on second scenario, using Kinect database 2, (20 participants, 25 records each).

In the third scenario the number of probabilities of selecting samples increased to ten, when the number of non-neutral gallery samples was increased by 1 for each of the 4 cases. The accuracy have increased, as a result of more representations of non-neutral cases in the gallery, by about 3%-5% across the different cases (see figure 6.8). These results are still below what was achieved by the last UGR scheme for the CASIA-B even when did not involve non-neutral samples in the Gallery (see Figure 5.4 and Table 5.4).

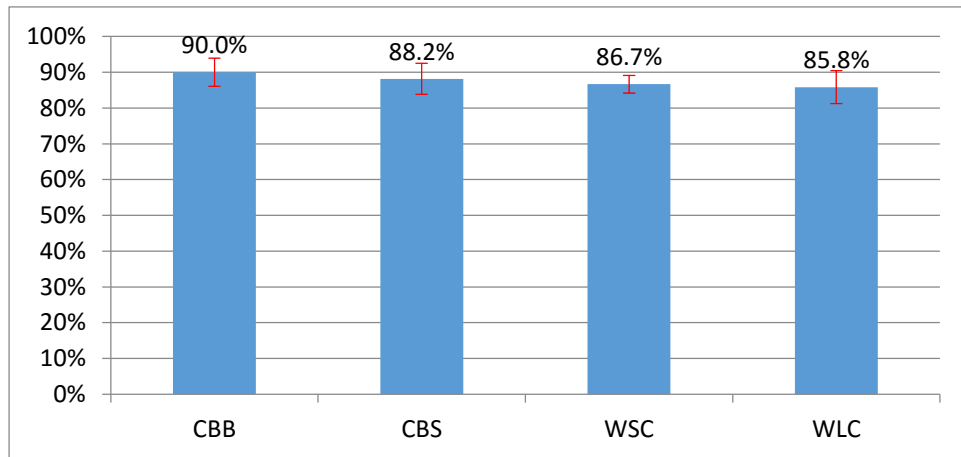


Figure 6-8: Recognition Performance (Average and standard deviation) of KUGR based on the third scenario using Kinect database 2, (20 participants, 25 records each).

The gallery in the fourth scenario includes 2 samples from each of the gait sequence variations including the neutral case. This arrangement allowed to testing of accuracy of the Kinect based scheme for the neutral case. Figure 6.9, show that the presence

of the balanced number of samples in the gallery has led to lower accuracy for both the neutral case, as was established in Figure 6.4, and for the WLC case. The reduction in accuracy for the neutral samples could be attributed to neutral walking could be confused with other non-neutral samples as a result of reducing the neutral sample in the gallery to 2. In the three other non-neutral cases this scenario has led to small improvement in accuracy.

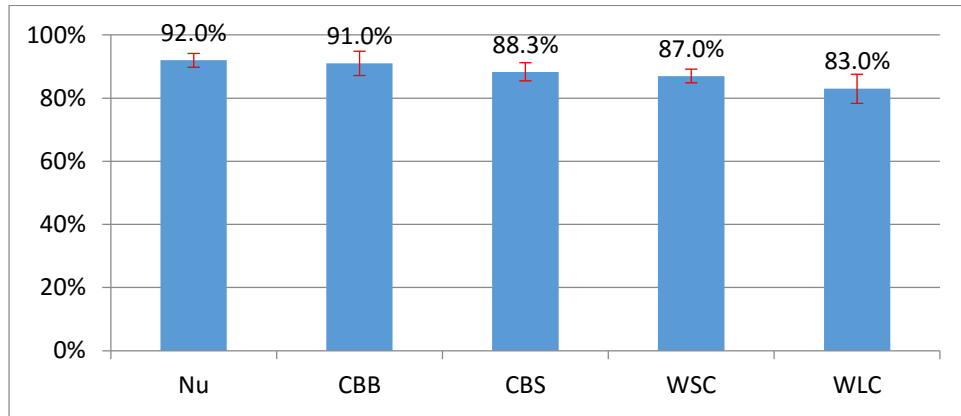


Figure 6-9: Recognition Performance (Average and standard deviation) of KUGR based on the fourth scenario using Kinect database 2, (20 participants, 25 records each).

In the last scenario the number of samples is reduced to be one for each type of the gait sequences. As expected, this has led to reduced accuracy compared to the last 3 scenarios for all cases. However, these results are significantly higher than the results that were achieved in the first scenario (see Figure 6.10).

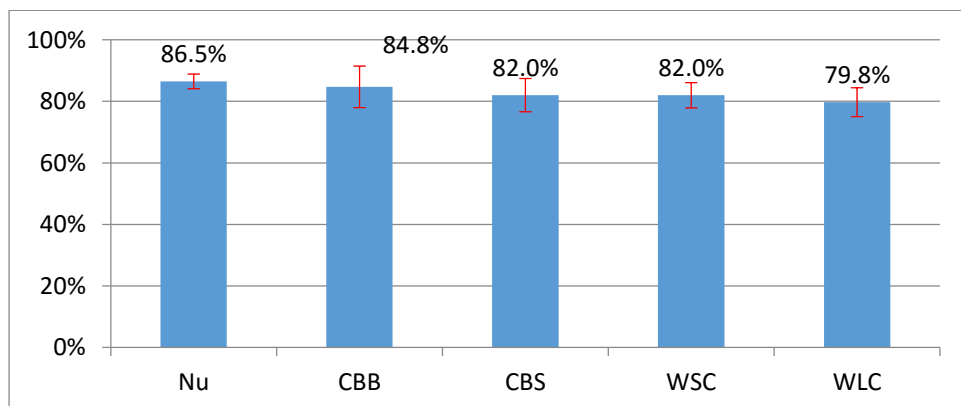


Figure 6-10: Recognition Performance (Average and standard deviation) of KUGR based on the fifth scenario using Kinect database 2, (20 participants, 25 records each).

6.4 Conclusions

In this chapter we presented human gait recognition approach based on skeleton point trajectories obtained from a Kinect sensor. We created two different database: the Kinect database-1 of neutral gait sequences only, and Kinect database-2 that includes both neutral and non- neutral gait sequences with balanced number of samples for each case. In our investigations we extended the number of Horizontal Distance Features (HDF) that are commonly used in the literature, and introduced the new set of Vertical Distance Features (VDF). We conducted gait recognition experiments using these features under different setup scenarios.

In the first experiments, which was conducted entirely on the neutral sequences, revealed that the KNGR scheme that is based on fused set of features (HDF+ VDF) outperformed the other single feature sets based schemes achieving 93% which is well below what was achieved by NGR in the last chapter on CASIA-B. Unlike the NGR, no gender recognition was integrated into KNGR. But this does not explain the differences in accuracy between NGR and KNGR, because KNGR confusing the gender was not noticed in its decision. Only in very few cases females (or males) were recognised as males (or females). In terms of classification method, the LDC classifier always outperformed the Naive Bayes method.

The performance of the unrestricted gait recognition scheme (KUGR) was tested under various protocols, but by only using the fused feature set (HDF+VDF) and the LDC classifier. The result that are presented in figures 6.6-6.10: shows that having more variation of the gait sequence in the gallery (scenarios 3, and 4) leads to better performance in the system. This is more evident in the first and fifth scenarios. The fifth scenario performed significantly better than the first scenario, due to having different gait sequence condition in the fifth scenario. The best results achieved was in scenario 4 with balanced samples per class were in the gallery, but these were still lower than what was achieved on the CASIA-B database in chapter 5.

In summary, the use of the Kinect sensor provide new opportunities for gait recognition, but more experiments, and bigger databases are needed to improve gait recognition specially in cases that have not been dealt with so far.

Chapter Seven

Conclusions and Future Research

7.1 Conclusions

In recent years the upsurge in terrorist attacks has focused research on biometric systems as a mean of identifying individuals from a distance or in a crowd. Gait is a behavioural biometric trait that is typically used to identify individuals while walking at a distance. Although gait has some advantages over other biometric traits for this purpose, the performance of gait recognition (GR) scheme are influenced by internal factors (pregnancy, physical disabilities, overweight, etc.) and external factors (carrying bags, wearing coats, view direction, types of footwear). This thesis was devoted to investigating human gait signatures, and develop gait recognition schemes that deals with neutral walking style as well as non-neutral cases characterised by the two most common external factors, carrying bags and/or wearing coats. Developing such gait recognition schemes that perform well in unrestricted gait sequences has shown to be a serious challenge.

Recognising a person from his/her gait sequence captured by a video camera require a pre-processing of the video frames detect the region of interest (the person's body or the smallest box around the body) and remove the background regions. Different schemes of background removal are adopted in the literature: like Frame-differencing and Mixture of Gaussian (MoG). These two algorithms have been implemented in spatial and in the wavelet domain, mostly in the low frequency LL-subband. The first is very efficient but has high error rates in outdoor scenarios, but the second performs well in- and out-door, but is time consuming. In this thesis, we first performed slightly enhanced versions of these two methods: In the first one, we used the average of first 10 frames, rather than the first frame, as a reference while in the second case we attempt to optimize the MoG scheme by dividing the frames into 5 equal size vertical blocks and use MoG on the first block. However, the last two algorithms are versions of the MoG using high frequency wavelet subband(s) rather than the spatial domain. First we calculate the standard deviation in the LL subband of the incoming frames, and only when this value increases beyond a threshold we trigger the MoG algorithm. In the first version, MoG applied to an integrated subband formed by choosing the

highest coefficient values in the 3 level 1 non-LL wavelet subbands of that image frame. In the second version the MoG is only applied to the LH subband alone. The advantages of these two procedures that the output is the body contour, and the last version is computationally the most efficient scheme. Although, the modified MoG procedures is successful in outdoor as well as indoor, in the rest of the thesis we only used our first algorithm for background removal due to the fact that current work is restricted to indoor gait recognition. However, the inclusion of the new wavelet-based MoG algorithm was originally for efficiency comparison purposes but it is motivated by our future plans to extend the current schemes to outdoor situation.

The first challenging step in developing a GR scheme, like in any biometric based recognition, post background removal is the extraction of quantitative features that can form a digital signature of gait cycles. Such feature are naturally dependent on the way the legs, hands and other body parts are changing during walk. Generally there two main types gait features: model-based (e.g. time series of geometric measurements of (or between) body parts over a gait cycle, or statistics of such measurements), and motion-based (e.g. gait energy image and gait entropy images). Motion-based features are also called the silhouette model. In the first component of this thesis, we developed various gait features that could be described as either, or even as a hybrid of the two models, but later for Kinect sensor investigations we extracted model-based features.

The various features developed in the literature are extracted mostly from the spatial domains of binarised frames but wavelet domain has also been used primarily the LL-subband as an approximation of the spatial domain. In this thesis we investigated a variety of gait-related features but have focused on extracting features from the wavelet domain, but additional to the use of the approximation subbands we exploit well-established statistical properties of the high frequency non-LL subband to provide a compact statistical model of motion during gait cycle. For the Kinect sensor investigation features were based on the skeleton points that are provided by the sensor and are determined in the spatial domain. The use of wavelet domain in this case can form part of future project.

For the first part of our investigations, features were all extracted from the wavelet domain either of the original video frames or from the Gait Energy/entropy images. We first defined three different sets of gait signature features were extracted from

different wavelet sub-bands at various levels: the STM (from LL1), the LMD (from LL1), and the AWM (from LL3). The STM expand the list of distance measurement features used in the literature and these features are all extracted automatically based on boundary dimensions calculated in separate parts of the body rather than entire body. The LMD has similar aim to the Energy image, but is different in that it consists of statistical features of inter-frame changes of binary pixel in the lower body part. AWM is based on the statistical parameters of LL3 subband coefficient for each single frame as well as the distribution of these parameters over the entire gait cycle. The use of LL3 was meant to remove possible effect of noise throughout the video frames. Each of the extracted features were subjected to the LDA procedure for dimension reduction and feature decorrelation, before using for gait recognition. The performance of each of resulting feature sets were singularly, and collectively by fusing them, investigated for neutral gait recognition (NGR). In total 7 different gait signature schemes were tested, and the experiments show that fusing the three vectors outperforms all other schemes with accuracy of 97%. This is comparable to results of the state of the art. PCA was also used but only when the three sets of feature were fused, without improving the accuracy compared to the LDA.

When repeated the best of the 7 NGR schemes with reduced number of samples in gallery, there was a drop in accuracy. On close examination we found that, for this reduced number of samples, most female misidentification were made with male and that this was possibly due to some features were more suitable for modelling male gait. This was a motivation to use some of the extracted features to develop a gender recognition scheme, to incorporate into GR. Beside GR, *Gender classification (GC)* is desirable in security surveillance, (recently we see frequently from the Middle East news that male terrorist tied to pretend to be female to get access through the check point easily). This was based on two sets of the same features used for NGR, but it was a 2-class problem. The k-NN classifier outperformed the SVM and achieved 96.5% GC accuracy. When incorporated the GC within the NGR, and used lower number samples in the gallery, the accuracy rate has improved modestly.

Unfortunately, when the best of the above NGR schemes was tested for recognising non-neutral gait sequences, the accuracy results were very disappointing, especially for the CW when the features were extracted from the whole body. We extended the AWM features, and created DWM feature vector to include similar statistical

parameters but from all high frequency subbands at levels 1, 2, and 3. This has led to 2 unrestricted gait recognition schemes (UGR), whereby in the first version fuses the NGR-based 3 feature sets with the DWM extracted from the whole body (upper and lower). In the second version we did the same except that AWM and DWM are extracted from the lower body part. The first version of UGR improved accuracy rates significantly for the CB case but by a small margin for the CW. The second version improved accuracy by large margins for the CW cases, but accuracy for Nu and CB were slightly lower than what was achieved the first version. By analysing results of UGR for females and males we found that in CW case, some of the misrecognised subjects (persons) were female and confused with male. We developed a GSCD method to detect the gait sequence case to be integrated into UGR. The experimental results showed that GSCD-UGR scheme provides significantly better result compared to UGR. Finally, incorporating gender classification with GSCD-UGR the result significantly improved by more than 4% achieving CB and CW are 99.6% (for Nu), 89.11% (for CB) and 85.48% (for CW).

We also investigated the use of the traditional features of Gait Energy Image (GEI) and the Gait Entropy Image (GENI) under unrestricted gait sequences. These two sets of feature are wavelet based rather than spatial domain and computed in different parts of the human body. Recognition based on these two schemes and their fusion resulted in high accuracy but not as high as those achieved by the GSCD-UGR. Combining the UGR and those based on GEI and GENI at the Decision level has led to the highest accuracy in comparison with the state of the art.

Finally we developed a gait recognition method based on extracting two different sets of features: Horizontal Distance Features (HDF) and Vertical Distance Features (VDF) from skeleton point trajectories, obtained from a Kinect sensor. This has provided opportunity to have balanced number neutral as well as non-neutral cases (Nu, CBB, CBS, WSC, and WLC). The experimental result on neutral was very successful. We also investigated the unrestricted gait recognition with the 5 case variations, and found that having different gait sequences cases in the gallery, will improve the performance of gait recognition.

From the above two settings, Normal cameras and Kinect sensors, we conclude that for high accuracy in the non-neutral cases, larger sets of features and/or involving samples of different cases are needed.

7.2 Future Work

Currently, a considerable amount of research relating to gait recognition is being developed, and the list of practical applications is growing fast. But, in comparisons to other biometrics, we need to overcome many hurdles before reaching the same level of deployment. Our experience throughout this project revealed a number of areas that require more investigations. Here we only list few promising directions for our future work.

1. **Expanding Kinect sensors work to outdoor setup.** In the beginning we shall compare the accuracy of the extracted skeleton points by Kinect sensor, with hand-labelled data. If we ensured that the points are extracted with good accuracy, then we expand the work on Kinect sensors. In this thesis we used the initial version of Kinect to extract HDF and VDF features from 20 body joints output by the sensor. The experiments demonstrated relatively good accuracy, but not for unrestricted gait sequence. In the new version the depth sensor has been enhanced and the tracking accuracy has been significantly improved. We plan to expand the database, by recruiting more participants with balanced gender representation, and more variety of unrestricted cases. Features, similar to those used for the CASIA B, will be extracted from the raw video recordings to be fused with HDF and VDF for unrestricted Gait recognition.
2. **Uncontrolled and outdoor Gait recognition.** Few research work, including the current thesis, deal with outdoor and uncontrolled situation, and yet it is the most relevant to real security and crime fighting applications. Future investigations will include investigations of gait recognition in uncontrolled environment to include variation in lighting conditions, variation in view angle and alignment, and outdoor recordings. As a first step, our efficient MoG background removal will be further developed to cope with such variation.
3. **Fusing Gait into Multi-modal biometrics.** This is primarily related to the challenge of identification of humans captured at a distance. We plan to investigate the fusion of gait and face biometric traits. These plans are linked to the above and focuses also to deal with outdoor scenarios.

References

- (PITO), P. I. O., (2005). *Biometrics Technology Roadmap for Person Identification within the Police Service*. In: PART 1: Identification Roadmap.
- Addison, D., Wermter, S. & Arevian, G., (2003). *A comparison of feature extraction and selection techniques*. Proceedings of International Conference on Artificial Neural Networks (Supplementary Proceedings, pp. 212-215.
- Aghaee, M., Amoon, M., Ayatollahi, A. & others, (2011). *Object Tracking in Video Sequence Using Background Modeling*. Australian Journal of Basic & Applied Sciences, 5(5).
- Ali, H., Dargham, J., Ali, C. & Mounq, E. G., (2011). *Gait recognition using gait energy image*. International Journal of Signal Processing, Image Processing and Pattern Recognition, 4(3), pp. 141-152.
- Al-Jawad, N., (2009). *Exploiting statistical properties of wavelet coefficients for image/video processing and analysis tasks*. In: Diss. University of Buckingham.
- AlMahafzah, H. & AlRwashdeh, M. Z., (2012). *A Survey of Multibiometric Systems*. arXiv preprint arXiv:1210.0829.
- <http://www.redbubble.com/people/coldblood/journal/5228103-composition-golden-ratio>.
- Arai, K. & Asmara, R., (2011). *Human Gait Gender Classification using 2D Discrete Wavelet Transforms Energy*. IJCSNS International Journal of Computer Science and Network Security, pp. 62-68.
- Arai, K. & Asmara, R. A., (2012). *Human gait gender classification in spatial and temporal reasoning*. Int J Adv Res Artif Intell, Volume 1, pp. 1-6.
- Arantes, M. & Gonzaga, A., (2011). *Human gait recognition using extraction and fusion of global motion features*. Multimedia tools and applications, 55(3), pp. 655-675.

- Ball, A., Rye, D., Ramos, F. & Velonaki, M., (2012). *Unsupervised clustering of people from 'skeleton' data*. In: *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. ACM, pp. 225-226.
- Bashir, K., Xiang, T. & Gong, S., (2009). *Gait recognition using gait entropy image*. In: *Crime Detection and Prevention (ICDP), 3rd International Conference on*. IET, pp. 1--6.
- Bashir, K., Xiang, T. & Gong, S., (2010). *Gait recognition without subject cooperation*. *Pattern Recognition Letters*, 31(13), pp. 2052-2060.
- BenAbdelkader, C., Cutler, R. & Davis, L., (2002). *Stride and cadence as a biometric in automatic person identification and verification*. *Automatic Face and Gesture Recognition. Proceedings. Fifth IEEE International Conference on*, pp. 372-377.
- Bharatkumar, A. et al., (1994). *Lower limb kinematics of human walking with the medial axis transformation*. *Motion of Non-Rigid and Articulated Objects*, , *Proceedings of the IEEE Workshop on*, pp. 70-76.
- Bharti, J. & Gupta, M., (2011). *Human Gait Recognition using All Pair Shortest Path*. *International Conference on Software and Computer Applications*, pp. 279-284.
- Bobde, M. S. S. & Satange, D., (2013). *Biometrics in Secure e-Transaction*. *International Journal of Emerging Trends & Technology in Computer Science*, 2(2).
- Bobick, A. F. & Johnson, A. Y., (2001). *Gait recognition using static, activity-specific parameters*. *Computer Vision and Pattern Recognition, CVPR 2001. Proceedings of the IEEE Computer Society Conference on*, pp. I-423.
- Borràs, R., Lapedriza, A. & Igual, L., (2012). *Depth information in human gait analysis: An experimental study on gender recognition*. In: *Image Analysis and Recognition*. Springer, pp. 98--105.
- Bouchrika, I., Goffredo, M., Carter, J. & Nixon, M., (2011). *On using gait in forensic biometrics*. *Journal of forensic sciences*, 56(4), pp. 882-889.
- Boyd, J. E. & Little, J. J., (2005). *Biometric gait recognition*. In: *Advanced Studies in Biometrics*. Springer, pp. 19-42.

Burges, C. J., (1998). *A tutorial on support vector machines for pattern recognition*. Data mining and knowledge discovery, 2(2), pp. 121--167.

Chang, C.-Y. & Wu, T.-H., (2010). *Using gait information for gender recognition*. In: Intelligent Systems Design and Applications (ISDA), 2010 10th International Conference on.IEEE, pp. 1388-1393.

Chang, S.-L., Hsu, C.-C., Lu, T.-C. & Wang, T.-H., (2007). *Human body tracking based on discrete wavelet transform*. Proceedings of the WSEAS International Conference on Circuits, Systems, Signal and Telecommunications, pp. 113-122.

Chen, L., Wang, Y. & Wang, Y., (2009). *Gender Classification Based on Fusion of Weighted Multi-View Gait Component Distance*. Pattern Recognition, CCPR . Chinese Conference on, pp. 1-5.

Cho, C.-W., Chao, W.-H., Lin, S.-H. & Chen, Y.-Y., (2009). *A vision-based analysis system for gait recognition in patients with Parkinson's disease*. Expert Systems with applications, 36(3), pp. 7033-7039.

Chourasiya, D. R., Richhariya, V. & Rajpoot, V., (2013). *Human Recognition Through Gait Energy Image And Principal Component Analysis Using Tumiitkp Database*. International Journal of Advance Research in Computer Science and Software Engineering.

Collins, R. T. et al., (2000). *A system for video surveillance and monitoring*. Carnegie Mellon University, the Robotics Institute Pittsburg.

Crnojevic, V., Antic, B. & Culibrk, D., (2009). *Optimal wavelet differencing method for robust motion detection*. Image Processing (ICIP), 16th IEEE International Conference on, pp. 645-648.

Cutting, J. E. & Kozlowski, L. T., (1977). *Recognizing friends by their walk: Gait perception without familiarity cues*. Bulletin of the psychonomic society, 9(5), pp. 353-356.

De Carrera, P. F., (2010). *Face recognition algorithms*. Master's thesis in Computer Science, Universidad Euskal Herriko.

Dikovski, B., Madjarov, G. & Gjorgjevikj, D., (2014). *Evaluation of different feature sets for gait recognition using skeletal data from Kinect*. Information and

Communication Technology, Electronics and Microelectronics (MIPRO), 37th International Convention on, pp. 1304-1308.

Duda, R. O. & Hart, P. E., (1973). *Pattern classification and scene analysis*. 3 ed. :Wiley New York.

Foster, J. P., Nixon, M. S. & Prugel-Bennett, A., (2001). *New area based metrics for gait recognition*. Audio-and Video-Based Biometric Person Authentication, pp. 312-318.

Gafurov, D., (2007). *A survey of biometric gait recognition: Approaches, security and challenges*. Annual Norwegian Computer Science Conference, pp. 19-21.

Gyaourova, A., Kamath, C. & Cheung, S., (2003). *Block matching for object tracking*. Lawrence livermore national laboratory.

Hafner, V. V. & Bachmann, F., (2008). *Human-humanoid walking gait recognition*. Humanoid Robots, 2008. Humanoids 8th IEEE-RAS International Conference on, pp. 598-602.

Han, J. & Bhanu, B., (2006). *Individual recognition using gait energy image*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 28(2), pp. 316-322.

Hassin, A.-A., Hanon, A., Talal, A.-A. & Barges, E., (2014). *Gait Recognition System Using Support Vector Machine and Neural Network*. Journal of Basrah Researches ((Sciences)) Vol, 40(4).

Heikkilä, J. & Silvén., O., (2004). *A real-time system for monitoring of cyclists and pedestrians*. Image and Vision Computing, 22(7), pp. 563-570.

Herman, I. P., 2007. Physics of the human body. *Springer Science & Business Media*. .

Hofmann, M. & Rigoll, G., (2012). *Improved gait recognition using gradient histogram energy image*. Image Processing (ICIP), 2012 19th IEEE International Conference on, pp. 1389-1392.

Hong, S., Lee, H., Nizami, I. F. & Kim, E., (2007). *A new gait representation for human identification: mass vector*. Industrial Electronics and Applications, 2007. ICIEA . 2nd IEEE Conference on, pp. 669-673.

- Hsia, C.-H., Chiang, J.-S., Dai, Y.-J. & Lin, T.-A., (2013). *Conditional-Sorting Local Binary Pattern Based on Gait Energy Image for Human Identification*. Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp),.
- Hu, M., Wang, Y., Zhang, Z. & Wang, Y., (2010). *Combining spatial and temporal information for gait based gender classification*. Pattern Recognition (ICPR), 2010 20th International Conference on, pp. 3679-3682.
- Hu, M. et al., (2013). *Incremental learning for video-based gait recognition with LBP flow*. Cybernetics, IEEE Transactions on, 43(1), pp. 77-89.
- Idrus, S. Z. S., Cherrier, E., Rosenberger, C. & Schwartzmann, J.-J., (2013). *A Review on Authentication Methods*. Australian Journal of Basic and Applied Sciences, Volume 7.5, pp. 95-107.
- Jain, A. K., Ross, A. & Prabhakar, S., (2004). *An introduction to biometric recognition*. Circuits and Systems for Video Technology, IEEE Transactions on 14.1, pp. 4-20.
- Jeevan, M., Jain, N., Hanmandlu, M. & Chetty, G., (2013). *Gait recognition based on gait pal and pal entropy image*. Image Processing (ICIP), 20th IEEE International Conference on, pp. 4195-4199.
- Jhapate, A. K. & Singh, J. P., 2011. *Gait Based Human Recognition System using Single Triangle*. International Journal of Computer Science and Technology, pp. 128-131.
- Jolliffe, I., (2002). *Principal component analysis*.:Wiley Online Library.
- Kamavisdar, P., Saluja, S. & Agrawal, S., (2007). *A survey on image classification approaches and techniques*. International Journal of Advanced Research in Computer and Communication Engineering, 2(1).
- Kameda, Y. & Minoh, M., (1996). *A human motion estimation method using 3-successive video frames*. International conference on virtual systems and multimedia, pp. 135-140.
- Katiyar, R., Pathak, D. & Kumar, V., (2010). *Clinical gait data analysis based on spatio-temporal features*. arXiv preprint arXiv:1003.1511.

- Katiyar, R., Pathak, V. K. & Arya, K. V., (2013). *A study on existing gait biometrics approaches and challenges*. International Journal of Computer Science, 10(1), pp. 135-144.
- Kim, D., Lee, S. & Paik, J., (2009). *Active shape model-based gait recognition using infrared images*. In: Signal Processing, Image Processing and Pattern Recognition. Springer, pp. 275-281.
- Kozlowski, L. T. & Cutting, J. E., (1977). *Recognizing the sex of a walker from a dynamic point-light display*. Perception & Psychophysics, 21(6), pp. 575-580.
- Kusakunniran, W., (2014). *Recognizing Gaits on Spatio-Temporal Feature Domain*. Information Forensics and Security, IEEE Transactions , pp. 1416-1423.
- Lee, B., Hong, S., Lee, H. & Kim, E., (2010). *Gait recognition system using decision-level fusion*. Industrial Electronics and Applications (ICIEA), 2010 the 5th IEEE Conference on, pp. 313-316.
- Lee, C.-C. et al., (2011). *Frame difference history image for gait recognition*. Machine Learning and Cybernetics (ICMLC), International Conference on, pp. 1785-1788.
- Lee, L. & Grimson, W. E. L., (2002). *Gait analysis for recognition and classification*. , Automatic Face and Gesture Recognition, Proceedings. Fifth IEEE International Conference on, pp. 148-155.
- Liu, T. & Wang, G., (2009). *A hierarchical approach for robust background subtraction based on two views*. Intelligent Systems. GCIS'09. WRI Global Congress on, pp. 325-329.
- Liu, Z. & Sarkar, S., (2006). *Improved gait recognition by gait dynamics normalization*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 28(6), pp. 863-876.
- Li, X. & Chen, Y., (2013). *Gait recognition based on structural gait energy image*. Journal of Computational Information Systems, 9(1), pp. 121-126.
- Lu, H., Plataniotis, K. N. & Venetsanopoulos, A. N., (2008). *A full-body layered deformable model for automatic model-based gait recognition*. EURASIP Journal on Advances in Signal Processing, Volume , p. 62.

- Mansour, R. F., (2012). *Multiple Views Effective for Gait Recognition Based on Contours*. Computer Engineering and Intelligent Systems, 3(10), pp. 50-62.
- Martín-Félez, R., Mollineda, R. A. & Sánchez., J. S., (2010) 'A'. *A gender recognition experiment on the CASIA gait database dealing with its imbalanced nature*. Int'l Conf. Computer Vision Theory and Applications (VISAPP), pp. 439-444.
- Martín-Félez, R., Mollineda, R. A. & Sánchez., J. S., (2011). *Human recognition based on gait poses*. In: *Pattern Recognition and Image Analysis*. Springer, pp. 347-354.
- Martin-Felez, R., Mollineda, R. A. & Sánchez, J. S., (2010) 'B'. *Towards a More Realistic Appearance-Based Gait Representation for Gender Recognition*. In: *Pattern Recognition (ICPR), 20th International Conference on*. IEEE, pp. 3810-3813.
- Mather, G. & Murdoch, L., (1994). *Gender discrimination in biological motion displays based on dynamic cues*. Proceedings of the Royal Society of London. Series B: Biological Sciences, 258(1353), pp. 273-279.
- Mayo, Z. & Tapamo, J. R., (2009). *Background subtraction survey for highway surveillance*. In: *Proc. of the Annual Symposium of the Pattern*.
- Meyer, D., (2004). *Support vector machines: The interface to libsvm in package E1071*. Citeseer.
- Migliore, D. A., Matteucci, M. & Naccari, M., (2006). *A revaluation of frame difference in fast and robust motion detection*. Proceedings of the 4th ACM international workshop on Video surveillance and sensor networks, pp. 215-218.
- Mir, A., Rubab, S. & Jhat, Z., (2011). *Biometrics verification: a literature survey*. International Journal of Computing and ICT Research, 5(2), pp. 67-80.
- Murray, M. P., (1967). *Gait as a total pattern of movement: including a bibliography on gait..* American Journal of Physical Medicine \& Rehabilitation, 46(1), pp. 290-333.
- Parajuli, M., Tran, D., Ma, W. & Sharma, D., (2012). *Senior health monitoring using Kinect*. Communications and Electronics (ICCE), 2012 Fourth International Conference on, pp. 309-312.

- Pratheepan, Y., Condell, J. & Prasad, G., (2009). *The use of dynamic and static characteristics of gait for individual identification*. Machine Vision and Image Processing Conference. IMVIP'09. 13th International, pp. 111-116.
- Preis, J., Kessel, M., Werner, M. & Linnhoff-Popien, C., (2012). *Gait recognition with kinect*. 1st International Workshop on Kinect in Pervasive Computing.
- Reynolds, D., (2002). *An overview of automatic speaker recognition*. Proceedings of the International Conference on Acoustics, Issue Speech and Signal Processing , ICASSP, pp. 4072-4075.
- Savage, R., Clarke, N. & Li, F., (2013). *Multimodal Biometric Surveillance using a Kinect Sensor*. Citeseer.
- Schuldt, C., Laptev, I. & Caputo, B., (2004). *Recognizing human actions: a local SVM approach*. *Pattern Recognition*, 2004. ICPR 2004. Proceedings of the 17th International Conference on, Volume 3, pp. 32--36.
- Sen-Ching, S. C. & Kamath, C., (2004). *Robust techniques for background subtraction in urban traffic video*. *Electronic Imaging* pp. 881-892.
- Sentayehu, E. W., (2006). *Face Recognition Using Artificial Neural Network..* AAU.
- Singh, J. P. & Jain, S., (2010). *Person identification based on gait using dynamic body parameters*. *Trendz in Information Sciences \& Computing (TISC)*, pp. 248-252.
- Sinha, A., Chakravarty, K. & Bhowmick, B., (2013). *Person identification using skeleton information from kinect*. ACHI 2013, The Sixth International Conference on Advances in Computer-Human Interactions, pp. 101-108.
- Sivapalan, S. et al., (2012). *The backfilled gei-a cross-capture modality gait feature for frontal and side-view gait recognition*. *Digital Image Computing Techniques and Applications (DICTA)*, International Conference on, pp. 1-8.
- Srivastava, H., (2013). *A Comparison Based Study on Biometrics for Human Recognition*. *IOSR Journal of Computer Engineering (IOSR-JCE)* Volume, Volume 15, pp. 22--29.

Stauffer, C. & Grimson, W. E. L., (1999). *Adaptive background mixture models for real-time tracking*. Computer Vision and Pattern Recognition, IEEE Computer Society Conference on..

Sudha, L. & Bhavani, R., (2011). *SVM based biometric authorization system by video analysis of human gait*. Electronics Computer Technology (ICECT), 3rd International Conference on, pp. 301-304.

T. Anitha, M. R., February (2013). *Gait Analysis for Gender Classification Using CASIA Gait Database*. International Journal of Innovative Research in Science, Engineering and Technology, 2(2), pp. 246-251.

Tang, J., Luo, J., Tjahjadi, T. & Gao, Y., (2014). *2.5 D multi-view gait recognition based on point cloud registration*. Sensors, 14(4), pp. 6124-6143.

Tao, G., Zhengguang, L. & Jun, Z., (2009). *Redundant discrete wavelet transforms based moving object recognition and tracking*. Systems Engineering and Electronics, Journal of, 20(5), pp. 1115-1123.

Tomasev, N., Brehar, R., Mladenic, D. & Nedevski, S., (2011). *The influence of hubness on nearest-neighbor methods in object recognition*. Intelligent Computer Communication and Processing (ICCP), IEEE International Conference on, pp. 367-374.

Töreyin, B. U., Cetin, A. E., Aksay, A. & Akhan, M. B., (2005). *Moving object detection in wavelet compressed video*. Signal Processing: Image Communication, 20(3), pp. 255-264.

Uhl, A. & Wild, P., (2008). *Footprint-based biometric verification*. Journal of Electronic Imaging, 17(1), pp. 011016--011016.

Venkat, I. & De Wilde, P., (2011). *Robust gait recognition by learning and exploiting sub-gait characteristics*. International Journal of Computer Vision, 91(1), pp. 7-23.

Wagg, D. K. & Nixon, M. S., (2004). *On automated model-based extraction and analysis of gait*. Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on, pp. 11-16.

- Wang, A.-H. & Liu, J.-W., (2007). *A gait recognition method based on positioning human body joints*. Wavelet Analysis and Pattern Recognition, 2007. ICWAPR'07. International Conference on, pp. 1067-1071.
- Wang, L., Ning, H., Tan, T. & Hu, W., (2004). *Fusion of static and dynamic body biometrics for gait recognition*. Circuits and Systems for Video Technology, IEEE Transactions on, 14(2), pp. 149--158.
- Wang, L., Tan, T., Hu, W. & Ning, H., (2003). *Automatic gait recognition based on statistical shape analysis*. Image Processing, IEEE Transactions on, 12(9), pp. 1120-1131.
- Wang, L., Tan, T., Ning, H. & Hu, W., (2003). *Silhouette analysis-based gait recognition for human identification*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 25(12), pp. 1505--1518.
- Wayman, J., Jain, A., Maltoni, D. & Maio, D., (2005). *An introduction to biometric authentication systems*. In: Biometric Systems. Springer, pp. 1-20.
- Xu, D., Huang, Y., Zeng, Z. & Xu, X., (2012). *Human gait recognition using patch distribution feature and locality-constrained group sparse representation*. Image Processing, IEEE Transactions on, 21(1), pp. 316-326.
- Yi, Z. & Liangzhong, F., (2010). *Moving object detection based on running average background and temporal difference*. Intelligent Systems and Knowledge Engineering (ISKE), International Conference on, pp. 270-272.
- Yogarajah, P., Condell, J. V. & Prasad, G., (2011). *P RW GEI: Poisson random walk based gait recognition*. Image and Signal Processing and Analysis (ISPA), 2011 7th International Symposium on, pp. 662-667.
- Yu, C.-C., Cheng, C.-H. & Fan, K.-C., (2014). *A gait classification system using optical flow features*. Journal of Information Science and Engineering, 30(1), pp. 179-193.
- Yu, S., Tan, D. & Tan., T., (2006). *A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition*. 18th International Conference on Pattern Recognition, Vol.4(IEEE), pp. 441- 444.

Yu, S. et al., (2009). *A study on gait-based gender classification*. Image Processing, IEEE Transactions on, 18(8), pp. 1905-1910.

Zamalloa, M. e. a., (2008). *Feature selection VS. feature transformation in reducing dimensionality for speaker recognition*. Proceedings of the V Jornadas en Tecnología del Habla, Volume 53.

Zhang, Z., (2012). *Microsoft kinect sensor and its effect*. MultiMedia, IEEE, 19(2), pp. 4-10.

Zhao, Z., Shen, Q. & Ren, F., (2013). *Heart Sound Biometric System Based on Marginal Spectrum Analysis*. Sensors, 13(2), pp. 2530-2551.

Zhou, J. & Hoang, J., (2005). *Real time robust human detection and tracking system*. Computer Vision and Pattern Recognition-Workshops. CVPR Workshops. IEEE Computer Society Conference on, pp. 149-149.